nature

17 June 1982

Cloud over French science plan

The French minister for research and technology's plan for future expansion may be delayed by the parliamentary process. But there is in any case a need to think about it more carefully.

The fall of the French franc on Sunday, which brings its devaluation against the German mark to something like 20 per cent in the past 12 months, may concentrate the attention of French parliamentarians wonderfully next week on a key point of the new research law, which is to be debated in the Assembly on Monday: its cost.

The *loi d'orientation et de programmation* for research and development is intended to be the principal instrument of the minister of state for research and technology, M. Jean-Pierre Chevènement. Almost everything hangs upon it: the restructuring of organizations such as the Centre National de la Recherche Scientifique, their stronger linkage with the economy, modifications in contracts of employment and — not least — the outline of a four-year budget. Under the terms of the bill, the government civil research and development budget, effectively Chevènement's budget, would grow in real terms by 17.8 per cent per year on the average as part of a strategy to bring total research and development expenditure in France to 2.5 per cent of Gross National Product by 1985. All basic research is guaranteed at least 13 per cent but high technological priorities (including biotechnology) would get increases of more than the 17.8 per cent. The absolute size of these annual increases is staggering — the best part of one National Science Foundation each year.

This ambitious plan was conceived in the early dreamier days of the present government. Even if the figures are agreed next week at the Assembly, it is doubtful whether they will be faithfully reflected in the 1983 budget, presently subject to horse-trading in the council of ministers. The problem is the overall French budget. The government planned a deficit of FF95,000 million in 1982, but the latest projections give a deficit of FF125,000 million largely because the French economy has not begun to grow at the expected rate of 3 per cent a year, so that tax returns are not covering expenditure. Moreover, increased social welfare benefits and the raising of the minimum wage have pushed up consumer spending but French industry has not been ready to respond, so that consumers have spent their money on imports instead. The government, alerted to the problem, has followed a well-trodden path in France and elsewhere in Western Europe and has imposed a classic squeeze: a four-month freeze on prices and incomes, combined with devaluation. So whither to grand plan now?

President François Mitterrand, speaking only a few days before the squeeze was announced — no doubt to prepare the ground — appeared to be phlegmatic. "We follow the same politics" he said. We are just in its second phase." His government's chief economic objective has been to reconquer the home market. In 1974, 24 per cent of French consumer spending went on imports, but the proportion had risen to 35 per cent by 1981. Part of the government's strategy has been priority for investment in nationalized industry (Mitterrand promised FF25,000 million in 1983) and for industrial innovation. But so far the President has refrained from mentioning a figure for research for 1983.

The research ministry is putting a brave face on this turn of events, but even if the government endorses its hopes for future budgets, trouble in the Assembly about other than financial matters could mean delay. The bill has in any case been badly mauled in the Senate, so that it will be for a joint Senate/Assembly committee eventually to pick up the pieces of Chevènement's grand plan — and to reassemble them as if into a jigsaw puzzle. If the job cannot be done before 1 July, the bill will have to wait until the winter.

The moral in this potentially disappointing tale is general and not specific. The present government of France was elected just a year ago on an optimistic programme of reform and renewal. Neither President Mitterrand nor M. Chevènement has exactly echoed Mr (now Sir) Harold Wilson's British pre-election speech of 1963 about the social benefits of "white-hot technology", but they have both been edging in that direction. The promise is valid, but the performance cannot be guaranteed. This is why the British experience in the 1960s should be studied in France. After the 1964 election, money was indeed thrown at a variety of technical projects but without much lasting effect. (An aluminium smelter built at that time in the Scottish Highlands was shut down only a few weeks ago, the predicted victim of the arithmetic of economics but a social catastrophe nonetheless.) In the end, the inflationary consequences of over-optimism precipitated the stagnation of the British science budget which has persisted ever since (Mr Edward Heath's intervening essay in what is now called Reaganomics notwithstanding). This is why other Europeans are distrustful of what the French have embarked on in the past year. The objective of spending more on research and development is entirely admirable, but nothing will be achieved if the bills have to be paid by printing money. For then it will be recognized that public expenditure on research and development is politically discretionary, and the tap will be turned off.

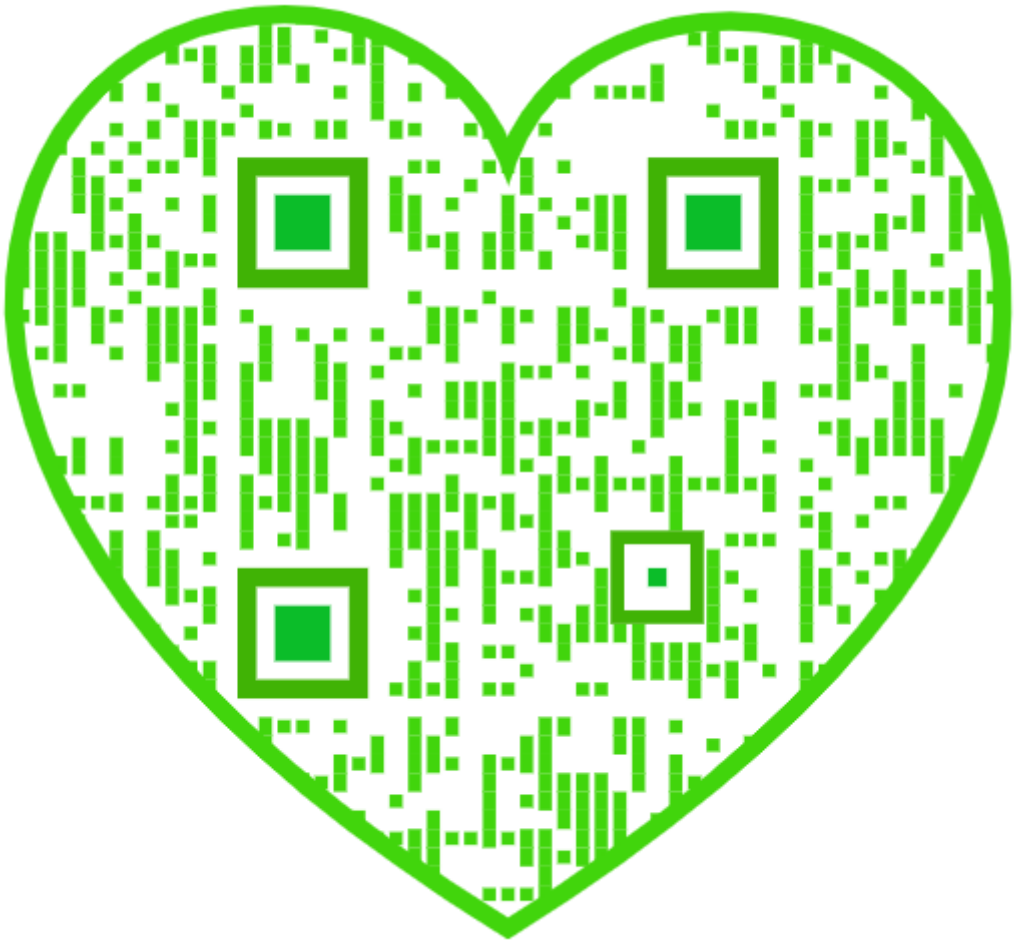
Who follows Slaughter?

Mr X, the next director of the US National Science Foundation, could spell trouble.

President Ronald Reagan has an unexpected opportunity for putting his own mark on the conduct of US science with the resignation of the director of the National Science Foundation (NSF), John B. Slaughter, whom President Carter appointed in 1980. Slaughter is ending his six-year term of office early, he says, only because he cannot resist an offer to become chancellor of the University of Maryland at College Park, the large state school with a high-quality science and engineering faculty. One of his aims is to integrate the university, which has always been an also-ran to its famous neighbour, Johns Hopkins University, more into the intellectual life of the Washington area. Slaughter says he also wants to attract top quality black students from the region to the university, instead of letting them be lured to other institutions elsewhere in the country. This ambition is understandable. Slaughter is one of the few black scientist-administrators in the United States, and one whose rapid rise has never been tainted with tokenism. (He was an assistant director of NSF, then vice-president and provost of the University of Washington before returning to NSF as director.) But another motive for his decision is the need to pay for his children's college education, and the Maryland job offers not only a pay rise but a free house. Similar reasons were given by another competent Administration figure, Admiral Bobby Inman, when he announced he would leave as deputy director of the Central Intelligence Agency, saying he would have to take a job in business. Must we conclude that the

0028-0836/82/240525-02$01.00

© 1982 Macmillan Journals Ltd



cost of college tuition is now so high that only business executives and university presidents can afford it?

Slaughter denies that being a Democratic appointee in the midst of the Republican Administration has caused problems, or that there is some particular disagreement behind his decision to leave. (The resignation is effective from 31 December.) This denial must be taken with a pinch of salt: the Administration has slashed NSF's science education programmes, particularly close to Slaughter's heart. But it is also true that he worked closely with the incoming Republican team to prevent NSF from being the target of serious cuts in the hard sciences. Is this why NSF has not been the butt of the broadsides of Dr Jay Keyworth, the President's science adviser, against the uselessness of the national laboratories or of planetary science?

In his brief term at NSF, Slaughter has made some important changes, along the lines of his early much publicized call to have NSF sponsor more applied research. He established a directorate of engineering, co-equal with the other directorates that run NSF's basic research. The new directorate, a combination of scattered programmes and some new things, now commands some 10 per cent of NSF's annual budget of approximately $1,000 million. Slaughter has and the National Science Board had a hard job persuading the basic research community to support the engineering directorate. At first, he says, they were "agnostic" about it (read "unenthusiastic"). Basic researchers in the hard sciences tend to regard NSF as their exclusive patron and historically have resisted its forays into public policy, engineering, social sciences and other matters. It is unfortunate, therefore, that Slaughter will not be around to follow through on the move. Past attempts to integrate engineering into the basic research agency have been conspicuously unsuccessful. Unless vigorously and carefully promoted, Slaughter's new directorate could yet founder. Presumably the job of promoting the engineering directorate will fall to Donald N. Langenberg, NSF deputy director, who had the job before Slaughter arrived.

Slaughter has also, with the help of the National Academy of Sciences, established a commission to develop a national policy on science and engineering education, especially urgent in the present crisis in technical education in US elementary and secondary schools (see *Nature* 6 May, p.9). Congress may yet give the commission funds to develop and implement a national plan. Slaughter says that the establishment of the commission is his single most important act as director of NSF. One of his merits is that these changes have been undertaken carefully, in consultation with NSF's traditional constituency, basic researchers in the hard sciences.

Mr X, the highly political candidate whom President Reagan might appoint to succeed Slaughter, could be a very different animal. Reagan is midway through his term of office, and is under fire from his original supporters, the extreme conservatives. If, to placate them, he appoints somebody (probably an industrialist) with fixed and preconceived views about the absolute merits of private enterprise, the result could be disaster. Mr X might well urge that NSF should get out of education policy and social sciences (both hotbeds of liberalism in the extreme conservative view) and no longer give seed money to small businesses doing basic research. NSF could finish up by being stripped down to its core mission — sponsoring basic research in universities — while its other more tentative but important roles, such as its recent brokerage in university–industry partnerships, vanish.

The effect of such a Mr X on the academic science community in the United States is easy to foretell. As a group, academic scientists tend to be paranoid about Washington's designs on their grant monies, and Mr X would strengthen their fears. The community would then pull back from its tentative experiments with industry, its unwilling acceptance, as Slaughter says, that "engineering is also important to the nation", and retreat to its ideological trenches. So it is to be hoped that Mr X will remain just what he is, a figment of the imagination, a hypothetical second-rater, whose *curriculum vitae* may be pressed on the President but whom the President will wisely decide not to appoint to the job Slaughter is vacating.

Nuclear plant and prices

The new chairman of the British electricity industry has a chance to make sense of the nuclear industry.

Dr Walter Marshall, chairman of the United Kingdom Atomic Energy Authority for the past year, is clearly once again the British government's favourite scientist. Last weekend he was made a knight, that peculiarly British public honour which requires that those whom he hardly knows should address him by his first name. A week before that, he was appointed chairman of the Central Electricity Generating Board, the principal owner of commercial nuclear power stations in the United Kingdom and the most likely source of further orders for nuclear plant. Much has plainly changed since 1978, when Marshall as chief scientist at the Department of Energy was for all practical purposes fired by his minister, Mr Tony Benn.

The most obvious change is the change of government (in 1979). But it is also clear that Marshall, whose reputation was previously that of a talented theoretical physicist who had succeeded by deftness and ebullience in carrying the Harwell establishment through hard times in the early 1970s, has used the interval to good effect. His most conspicuous success in the past year has been to persuade the engineers who will be working for

him from 1 July that it is possible to build pressurized-water reactors in Britain at less than twice the cost that obtains elsewhere. The absorbing question now is what Marshall's arrival at the generating board will mean for the British nuclear industry.

The precedents are not uniformly encouraging. Sir Christopher (now Lord) Hinton, who built the first dozen or so nuclear reactors in Britain as well as the separation plant at Windscale (now called Sellafield) maddened his ex-colleagues when, as chairman of the generating board, he insisted that the only significant difference between nuclear electricity and other kinds of electricity is the relative cost. Logically, that view is unshakeable, but it does not thereby follow that the customers (especially monopoly customers) for nuclear plant are without influence on the costs that they must pay. For too much of the past two decades, the generating board has been too passive a customer for nuclear plant, looking to the government to be told when to build reactors, putting up compliantly with the publicly created suppliers' monopoly (called the National Nuclear Corporation) and, more generally, behaving limply.

Marshall, fortunately, is far from limp. The most likely immediate consequence of his translation to the generating board is that this huge nationalized industry will become less dull. The chance of putting the British nuclear industry on a sound commercial footing — or at least of telling conclusively whether such a goal is feasible — is higher now than ever. But in the end what matters is that there should be an asssured supply of electricity at the lowest possible cost. Nuclear power undoubtedly has an important part to play, but so too does the elimination of high cost (oil-fired) and excess capacity from the board's network of generating plant. Pure reason, alas, is not sufficient. Marshall's imaginativeness will help. So too will his capacity to make even those whom he has talked down think well of him.

0028-0836/82/240526-01$01.00
© 1982 Macmillan Journals Ltd

Political interference at FDA?

New policy raises fears of meddling

Washington

For the second time in as many months, the Reagan Administration is being accused of injecting politics into the selection of scientists for independent advisory panels. According to an official at the Food and Drug Administration (FDA), appointment authority over its scientific advisory panels is gradually being taken away from the FDA commissioner and given to the Department of Health and Human Services (HHS). What worries FDA officials is that HHS recently suggested at least two scientists whom FDA believes are unqualified and who were proposed for political reasons.

Earlier this year, the Department of Agriculture admitted that it had been running security checks and "political" checks on scientists nominated for peer-review panels to evaluate proposals for its competitive grants programme but the political checks have since been dropped.

The troubles at FDA, however, seem more serious. Although the secretary of HHS has always had authority to appoint the advisory panels at FDA, that function has traditionally been delegated to FDA itself. One FDA official said that although HHS has given general directions on the numerical balance of the committees; approval by HHS was generally a formality.

Over the past several months, however, HHS has rewritten the panels' charters as they come up for renewal, retracting FDA's delegated control over them. Although FDA is still being asked to identify possible candidates it is also, for the first time, being told to furnish a list of alternatives. And the final decisions are apparently now being made in the office of HHS secretary Richard S. Schweiker.

Of even greater concern to FDA is that the secretary's office has begun to offer suggestions of its own. According to a report in the *Washington Post* last week, which was confirmed by FDA, these suggestions included a woman psychiatrist who was a founder of the "California Pro-Life Council," an anti-abortion group. She was suggested by HHS as the consumer representative on a panel that evaluates contraceptive and abortion drugs. FDA told the secretary's office that she was unqualified, but whether FDA's opinion will be considered is still uncertain.

The *Washington Post* also reported that HHS suggested a California physician whose main qualification, according to FDA scientists, was apparently that he listed Dr Loyal Davis — Nancy Reagan's stepfather — as a reference.

The Department of Health and Human Services claims that it is doing nothing improper. An HHS official said the department is merely continuing the centralization of appointment authority begun under President Carter. The official also denied that the secretary's office was making its own recommendations for committee positions, and said it was simply passing on to FDA all names that it receives from congressmen and senators of both parties.

That interpretation was contradicted by an FDA official, who said that while names had occasionally been passed on before, what was happening now was quite different. The official said that HHS was coming back with new names *after* FDA had submitted its list of candidates, and was specifically asking if these new names were qualified for specific positions. Mr Garret Cuneo, the HHS official now apparently in charge of the selection process, refused to comment on the matter.

The effects of the new appointment policy remain to be seen; no obviously political appointments have been made so far. But even if political manipulation does not occur, FDA is worried that the new procedures will make putting together an advisory panel even harder. The aim in the past has been to get not only leading scientists in a range of specialities, but also to have different geographical regions and women and minorities represented on each panel. FDA officials wonder how these goals can still be met when it is cut out of the process and when it is limited to saying whether the HHS candidates are qualified or not.

Meanwhile, at the Department of Agriculture, the dust seems to have settled. According to the office of Agriculture Secretary John R. Block, the decision to run political checks on scientists nominated for peer-review panels was really the result of a misunderstanding. Political checks are normally run on policy advisers; last autumn, the Office of Management and Budget ordered the department to bring the peer-review panels under the same federal act that governs policy panels. Agriculture department officials then assumed that political checks would have to be run for them as well.

After disclosure of the political checks on scientists, Block ordered them to be stopped. The scientists who had not yet been cleared by Block's office were then all approved at once. Block's office maintains, however, that none of the information

NIH urged to act on germ war

Washington

In response to rumour that the US Army and Navy are seeking to develop "defensive biological weapons systems" using recombinant DNA techniques, two American biomedical scientists have called on the National Institutes of Health (NIH) specifically to prohibit such work.

Dr Richard Goldstein of Harvard Medical School and Dr Richard Novick of the Public Health Research Institute of New York City are proposing an amendment to the NIH recombinant DNA guidelines to ban "the construction of biological weapons by molecular cloning". The proposal will be considered at the next meeting of the Recombinant DNA Advisory Committee on 28 June.

According to Dr Goldstein, several scientists have recently been approached by the military about the possibility of using cloning techniques to produce biological warfare agents. The United States signed the 1982 Biological Weapons Convention, agreeing "never in any circumstances to develop, produce [or] stockpile" biological agents for other than peaceful purposes. But Dr Goldstein says the military apparently interprets this as not covering the development in the laboratory of suitable organisms for "defensive" — by which is meant deterrent — purposes.

In a statement describing their proposed amendment, the scientists argue that "the use of molecular cloning for the deliberate construction of biological weapons is, *per se*, the most serious biohazard imaginable for this technology", and that "it constitutes an egregious misuse of scientific knowledge".

Although the NIH guidelines strictly apply only to researchers working under NIH grants, the Defense Department has so far agreed to follow them. But the amendment's sponsors say that even if their proposed ban is not binding on military research, it would "provide automatic public support for a refusal of the scientific community to participate in the development of biological weapons and it may convince governments that the 1982 prohibition should be construed as applying to laboratory research".

Under the recent relaxation of the DNA guidelines (see *Nature* 29 April, p.793), the previous ban on cloning of toxin genes and on release of recombinant organisms into the environment was lifted, although permission from NIH is required before proceeding with either. The amendment would restore the ban in these two areas when the aim is construction of biological warfare agents, and extend it to previously unregulated activities such as the cloning of viruses for this purpose.

Stephen Budiansky

0028-0836/82/240527-02$01.00

© 1982 Macmillan Journals Ltd

obtained in the political checks was used.

The department will continue to run nominees' names through an FBI security check which it calls "routine", although neither the National Institutes of Health nor the National Science Foundation follows this practice.

Professor Joe Key of the University of Georgia, a director of the competitive grants programme under President Carter, wonders if the checks were an attempt to cause trouble for the four-year-old programme, which is still small and fragile and viewed with some suspicion elsewhere in the department. The bulk of agricultural research (some $1,500 million a year, of which $15 million is for competitive grants) is funded through land-grant colleges and state experimental stations, which receive government funds according to an automatic allocation formula. This latest incident did little to help the competitive programme's vulnerable position.

Stephen Budiansky

West German environmentalists

Greens' delight

Heidelberg

The Hamburg elections on 6 June have left the environmentalist "green" party in a powerful position. While the conservative Christian Democratic Union (CDU) won one more seat in the Hamburg Senate than the Social Democrats (SPD), the greens (Grün-Alternativ-Liste) won nine seats and now hold the balance of power.

In Hamburg, CDU leader Walther Leisler Kiet is now trying to form a non-party "citizens' senate". If he fails, Mayor Klaus von Dohnanyi of SPD will have to ask the greens for their support. The greens have set their faces against coalition and have said that they will cooperate only on strict conditions, including Hamburg's abandonment of nuclear power.

The aftershocks of the Hamburg election continue to reverberate in all political strata in West Germany. The personal intervention of Chancellor Helmut Schmidt and Foreign Minister Hans Dietrich Genscher failed to turn the tide in favour of the Social Democrats, while the Free Democrats, the lesser partner with the Social Democrats in the Bonn coalition, again failed narrowly to reach the 5 per cent threshold necessary to qualify for a seat.

The polarized vote in Hamburg reflects the shakiness of the Bonn coalition as well as the strength of the green-left. If the Social Democrats lose only one more *Land*, the Christian Democrats will have a clear two-thirds majority in the Bundesrat and could block all government legislation. Rumours are now rife that the Free Democrats will leave the coalition after the Hesse election on 26 September, perhaps joining with the Christian Democrats. The deepening green intrusion into West German politics is at least partly responsible.

With seats in five parliaments, the greens have become the third strongest party. The abrasive challenge they offer to the assumptions of the established social industrial system is proving hard for even the Social Democrats to meet. They are concerned that "integrating" green issues would further alienate the traditional working class base without necessarily keeping ecologically-minded young voters. Indeed, there are wide differences between the Social Democrats and the greens except on individual issues such as the Rhine-Danube canal or the protection of wildlife.

The central goals of the greens are the realization, by means of a social welfare state, of the recommendations of Global 2000 and of disarmament. They call, for example, for radical change to a no-growth society, the use of alternative technologies and energy sources, conservation, economy and peace. They oppose the "compromise politics" of the Social Democrats.

Voters are mainly young and well educated and the focus is often local: cauliflower cancers of eels from the polluted Elbe and the discovery of tonnes of nerve gas in a ramshackle chemical factory near the city centre were important at Hamburg. The parties are autonomous at every level, so that there are no declared leaders and the small new national party controls neither the policies nor the attitudes towards coalition of the *Länder* parties. To ensure the five per cent vote, the Länder-Grünen have often joined with other groups, some of whom have made uncomfortable allies.

Their anarchic demands for Rousseau-esque self-determination, their use of violence and frequent threats to make West Germany ungovernable if minority demands are not met raise spectres of the Weimar Republic.

In September, the Hessen greens will go it alone, campaigning especially against the extension of Frankfurt Airport and plans for four local reprocessing plants for atomic waste. They expect a 10 per cent vote, a Christian Democrat parliament and a defeat for the Social/Free Democrat alliance that would indeed have national repercussions.

Sarah Tooze

European science policy

Give and take

Brussels

The second report on European science policy by Crest, the Committee on Scientific and Technical Research, tells the familiar story of how, after 15 years of trying to bridge the technology gap with the United States, Europe still lags behind in the latest technologies. Of the million or more European scientists — representing 20 per cent of the world's total — 350,000 are researchers, and science in Europe cost $35,000 million in 1980. The new report, which has still to be published, goes beyond its predecessor in comparing the national research and development policies of the ten EEC member states.

Energy research, which since 1974 has been a priority, accounted for a smaller share of the research and development budget at 11 per cent in 1980 than in either the United States (12 per cent) or Japan (25 per cent). Nuclear (fission) research has seen its importance slump, particularly in West Germany and France.

Allocations to industry research grew by 15 per cent a year in the Community as a whole between 1978 and 1980 but are thought to have returned to pre-1974 levels only last year. The increase has been marked in Italy and the United Kingdom, but UK spending is still below the 1974 level.

Space research, where international cooperation is at its most developed in Europe, shows France and West Germany sharing more than two-thirds of the spending, worth $860 million in 1980. Curiously Italy and Belgium, recent years have contributed the greatest proportion of their budgets to space research.

Agricultural research turns out to play the biggest role in the smaller member states. In Ireland and Greece a quarter of public research and development money goes on agriculture, which is equivalent to only 3.7 per cent of the EEC's total spending. Agriculture is also important in Denmark and the Netherlands which spent $3.7 and $6.8 per head respectively compared with an EEC average of $2.7, and $3.1 in the United Kingdom. Italy's outlay seems small at $0.9 per inhabitant, considering agriculture's social and economic importance in the country.

International cooperation is still important but its share of public budgets has fallen by as much as 25 per cent since 1978. Between 1974 and 1978 it was constant at 10 per cent but now there are wide differences among the member states. Britain's contribution, for example, has increased by 25 per cent since 1970.

The ten members of the EEC also differ in the response of their companies to the technological challenge. The role of the state in industrial research is greatest in the United Kingdom (30 per cent), France (22 per cent) and West Germany (20 per cent), but less than 10 per cent in the rest. Measuring research in companies as the proportion of expenditure on research and development to the value added of the manufacturing puts British companies first. The proportion of research expenditure is 4.7 per cent followed by West German companies (3.9 per cent), French companies (3.5 per cent) and Dutch companies (3.3 per cent); in the United States though, the average is 5.7 per cent. But if a distinction between public expenditure and companies' own funds is made, the picture changes. US companies spend the most (3.8 per cent) followed by the Japanese (3.4 per cent).

The report argues that many of the

0028-0836/82/240528-02$01.00

© 1982 Macmillan Journals Ltd

member states are not pulling their weight. Germany, France and the United Kingdom share 80 per cent of Community research and development expenditure while the combined efforts of Belgium, Denmark, Ireland and Greece amount to as little as 6 per cent. The conclusions are being studied in preparation for the next EEC science council of ministers on 30 June.

Jasper Becker

Luxembourg in space

TV plan delay

Luxembourg is falling behind in the race to provide the first satellite broadcast for Europe. Unresolved wrangles about the relative sizes of French, Belgian and West German shareholdings in the projected £200 million Luxembourg system have now put back the earliest possible launch date until late 1986.

By then, the national satellites planned jointly by France and West Germany will probably have been up for a year. Also by late 1986, the new all-British Unisat, with two BBC channels, might be aloft.

Radio Tele-Luxembourg (RTL), the commercial company in charge of all broadcasting in Luxembourg, insists that its satellite project is a matter of when, not if. It plans to beam three channels of films, news, entertainment and advertising to a wide swathe from Benelux to Bavaria, and possibly to south-east England. It knows, however, that this ambition represents a big gamble for a small country. RTL, whose current radio and television programmes reach a weekly audience of 40 million in Europe, contributes 5 per cent (about £25 million) of the annual income of the Luxembourg government. Commercial broadcasting has been a major source of revenue for the 50 years since Luxembourg, accustomed to living on its wits, recognized that its central location, plus national sovereignty over its broadcasting policy, permitted it to send radio advertisements and popular music into countries whose governments prohibited one or both by their own broadcasting organizations. It was in 1932 that Radio Luxembourg began spoiling the determination of Lord Reith, the BBC's first director-general, that the British people should have no light radio entertainment on Sunday.

Now, however, with commercial radio and television sprouting all over Europe and eroding its audience, RTL must find a new service to sell or see its profits and contribution to the national purse atrophy. But even if its satellite succeeds in pulling in audiences and advertising revenue, the heavy costs of depreciation of the start-up years is going to hurt Luxembourg's economy, already in trouble with its first unemployment since the Second World War.

RTL cannot proceed, however, until it decides upon a new financial structure to accommodate the necessary infusion of new capital. Although the company is registered in Luxembourg and has, by law, a majority of Luxembourg nationals on its board, its largest shareholder is predominantly Belgian. Also, French investors, notably in the form of Havas-IP, Compteurs Schlumberger and Paribas, are in a strong if not controlling position. RTL's problem now is how to introduce a

big new outside investor — as a group of West German publishers (which does not include the giants) has offered about £50 million — without disrupting the delicate Franco-Belgian balance of power. The French hope to avoid excessive dilution of their own influence. But they, and the whole board, know that the matter will have to be sorted out next month if further delay is not to occur.

Once the go-ahead is given, RTL will invite tenders for the system (two satellites plus a spare). It will also decide whether to devote one of the three television channels to broadcasting in English. Any deregulatory move by the Hunt committee now inquiring into the expansion of cable television in Britain would undoubtedly sway the decision. Elsewhere, RTL insists, a vast expansion of cable television in Europe is not necessary to the success of its plan. It expects half of its audience to buy a rooftop dish to receive its programmes.

Brenda Maddox

Monsanto hands out $23.5 million

St Louis

The Monsanto Company has awarded $23.5 million to the medical school at Washington University in St Louis, Missouri, for five years of research into how proteins and peptides affect cell regulation. The grant, announced on 3 June, approximately doubles the research funds from non-federal sources at the medical school for that period, and is one of the largest from a single company to a university.

Monsanto's award is an obvious boon to the university's research programme: two-thirds of the money will go to applied research and one-third to basic research. Work will be selected by an eight-member committee — four university faculty members and four people from Monsanto — headed by Dr David M. Kipnis, head of the department of internal medicine at the university medical school. University researchers will be at liberty to publish the results of work funded under the grant, but if the material contains potentially patentable technical developments Monsanto can review it, and request a short delay before submission for publication. Monsanto will have exclusive rights to any licences arising from patents on the work, but patents will be the exclusive property of the university, which will be able to receive royalties from Monsanto licences. Royalties will go to the university's research and education programmes — not to individual researchers.

Meanwhile, as many as "a couple of dozen" Monsanto scientists may work at the university and some may spend several years there. Monsanto's move clearly shows an interest in medical products, and Dr Howard A. Schneiderman, Monsanto's senior vice-president for research and development, says that while the company does not now market health-care products, it hopes to do so in future.

Monsanto launched a $1.8 million project with Washington University earlier this year for research using monoclonal antibodies for diagnostic work. The company is also working with Dr Mary-Dell Chilton, of the university biology department, on genetic engineering in plants. Schneiderman did not, however, confirm rumours that a major new grant from Monsanto was imminent to support Chilton's research.

Karen Freeman

Jobs and automation

US faces facts

Washington

The effect that new electronic technology is having on jobs, which has been an issue of intense debate in Europe, is now a cause for concern in the United States.

A new study by the government's General Accounting Office (GAO) suggests that the recent revolution in electronics will be felt not just in manufacturing — the sector most affected so far by automation — but also in office and service jobs. And even within manufacturing, the spectrum of fixed automation introduced in the 1950s, and the initial applications of robots lately have tended to be in tasks that were considered menial, or monotonous or unsafe, for human workers. Spray painting

0028-0836/82/240529-02$01.00

© 1982 Macmillan Journals Ltd

of automobile bodies is an example. Now, with microprocessors that can be incorporated into virtually every machine, even skilled and previously immune occupations such as tool-and-die making may be affected.

The critical question is the validity of what has been a traditional assumption: that technological progress brings with it more jobs. Several witnesses scheduled to testify next week before a House subcommittee investigating the issue doubt that assumption still holds. "We can't count on expansions in the white-collar or service areas, which is what saved us in the fifties and sixties", says William Bittle of the International Association of Machinists and Aerospace Workers.

An annual employment forecast issued by the Bureau of Labor Statistics (BLS)* confirms that employment in at least some office and service occupations is being hit by electronic technology. "We do see some jobs disappearing", says Ronald Kutscher, assistant commissioner of BLS for economic growth and employment projections. Key-punch operators, telephone operators and virtually all the printing trades will be hard-hit, for example.

The other open question, though, is how many jobs will be created by the new technologies directly. Workers will be needed to build, install, adjust, and repair automated equipment. The GAO study found virtually no evidence that could answer this question, however.

The trade unions have apparently accepted that jobs will be displaced by automation. But the critical issue to them is whether enough time will be allowed for workers to find new positions. In Norway, unions have negotiated contracts that set a gradual rate for the introduction of new technologies. The possibilities of such contracts being agreed to in the United States seem much smaller. A common complaint by American trade unions is the tendency towards secrecy on the part of management and the absence of the sort of cooperation and consultation practised in Europe and Japan.

The House subcommittee hearings may be a small step towards some government action on the problem. Representative George Miller, who is holding the hearings, has introduced a bill (HR 5820) that would provide for vocational retraining of displaced workers in new occupations created by automation. The unions, however, tend to dismiss government-supported training as a subsidy for industry and an inefficient substitute for on-the-job training. More to the point may be another concern of Miller's: he points out that the government spends nearly $2,000 million a year on labour-saving devices. **Stephen Budiansky**

Advances in Automation Prompt Concern Over Increased U.S. Unemployment (General Accounting Office, May 25, 1982). *Occupational Outlook Handbook* (U.S. Department of Labor Bureau of Labor Statistics, April 1982).

GM cancer prizes

Rules to be bent

Although this year's General Motors Cancer Research Prizes have been duly awarded (see below), leaving Dr Howard Skipper, Dr Denis Burkitt and Dr Stanley Cohen each $100,000 better off, the awards committees are clearly running into difficulties in selecting an annual trio of winners while sticking to the rules. Only four years after the awards started, the biggest worry is that of finding each year someone worthy of the prize "for the most outstanding recent contribution to the prevention of cancer, including environmental factors".

The rules of the prizes were set in 1978 when General Motors, disturbed by the number of its directors who had become victims of cancer, put $2 million (just doubled) into a General Motors Cancer Research Foundation. The prizes are large enough to invite comparison with the Nobel awards; the rules, however, differ in interesting ways.

One rule, intended to eliminate fortuitousness, is that a prize winner should have made more than one major discovery. Their discoveries must have been made within the previous fifteen years unless their importance has been recognized only more recently.

One prize (Kettering) is for diagnosis and treatment of cancer, another (Mott) for prevention and the third (Sloan) for a contribution to basic science. Winners are chosen by a process that resembles that used for the Nobel prizes. From a list of 25,000 prominent scientists, about 6,000 each year are asked to nominate candidates. Three subcommittees, one for each prize, first pare the nominations to twelve. Last year, they had to sift through 114, 40 and 91 nominations respectively. By the second meeting, each committee member has to report on the merits of two of the twelve candidates, eight of whom are then eliminated. At a final meeting the committees rank two of the four remaining candidates in order of preference. Finally the awards assembly has to decide whether to follow the committee's advice.

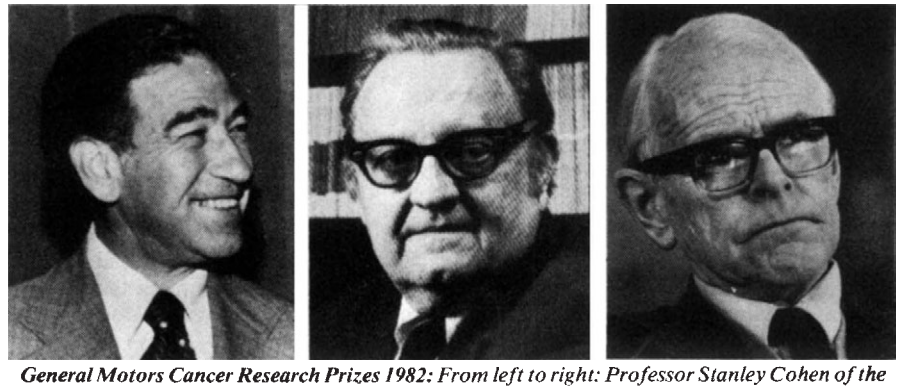
This year the assembly argued whether it should award the Mott prize for prevention. Nobody seems to have doubted the importance of Dr Burkitt's discovery of the childhood cancer that now bears his name (Burkitt's lymphoma) and his perceptive suggestion that it is transmissible (it later became clear that a virus is involved). Nor is it in doubt that he pioneered the chemotherapy of "his" lymphoma. But that was all more than fifteen years ago and in any case cannot strictly be considered a contribution to the prevention of cancer.

Turning a blind eye to those problems, the relevant committee and the assembly also had to grapple with the question whether Dr Burkitt's advocacy of the importance of dietary fibre in the prevention of cancer, the topic that has most occupied him in the past fifteen years, is more than a provocative hypothesis. In the end, it was not taken into account.

The choice of Dr Howard Skipper for the Kettering prize for diagnosis and treatment ran into much less opposition, although again the rules have obviously been stretched. Skipper is widely acknowledged as a pioneer of cancer chemotherapy. For 35 years he has influenced clinical chemotherapy by extensive studies on animal and cell models. His discoveries have influenced which drugs are used, in what combinations and their dosage and timing. It is, however, not easy to point to two major discoveries of Skipper's within the past fifteen years. His most recent work bears on the understanding of drug resistance in tumours.

Even for the least disputed of this year's prizes — that to Dr Stanley Cohen — an elastic interpretation of the rules is evident. There is no doubt that he put epidermal growth factor on the map and that it is relevant to cancer research. Cohen's earlier and very important work on nerve growth factor is not a contribution to cancer, and so the characterization and the biological effects of epidermal growth factor have had to be considered separate discoveries.

Perhaps prize rules are made to be stretched. Certainly as Robert Burton once said: "No rule is so general, which admits not some exception". But when the exception is the rule it may be time to change them. **Peter Newmark**



General Motors Cancer Research Prizes 1982: *From left to right: Professor Stanley Cohen of the department of biochemistry at Vanderbilt University School of Medicine; Dr Howard E. Skipper, recently retired president of Southern Research Institute, Birmingham, Alabama; and Dr Denis Burkitt, honorary senior research fellow, St Thomas's Hospital, London.*

0028-0836/82/240530-01$01.00

© 1982 Macmillan Journals Ltd

Awaiting the gypsy moth

Washington

Spring in New England, with its green-shaded campuses, tall shadowy and seemingly endless deciduous forests, and country villages with white church spires set against rolling green mountains, used to be one of the natural glories of North America. But as spring advances in New England this year, so will the gypsy moth (*Lymantria dispar*). During June and July the moths will emerge from their pupae and begin their destructive predatory march that now covers most of eleven states from Maine down to Maryland, munching away at the entire foliage of white, red, black and scarlet oaks, and in infested areas of the grey birch trees, the apple trees, American beeches, red maples and even conifers including the common white pine. In areas of heavy infestation, the upper ridges of mountains become so entirely denuded of tree canopy and ground cover that they are ashen and apparently lifeless, reminiscent, as one New Haven scientist said last week, of an experiment that Brookhaven National Laboratory ran in the 1950s to see what a forest would look like after irradiation by nuclear weapons.

The plague of the gypsy moth in the United States is a genetic experiment run wild. In 1869, L. Trouvelot, a French astronomer and naturalist who lived at 27 Myrtle Street, Medford, Massachusetts, imported some gypsy moths from Europe to help found a silkworm industry in New England, then the centre of the textile industry. Some of the eggs were lost, and by 1889 they had multiplied enough to provoke an outbreak in the town of Medford. The following year the state legislature awarded $25,000 for control.

Things have got worse since. The larvae hang on silk threads they have spun, so that air currents can carry them to new trees in which to hatch. Prevailing winds encouraged the spread through northern New England. In 1923 the government established the north–south flowing Hudson River as a natural "barrier zone" to contain the spread south. The Second World War interrupted control efforts, but brought an apparent remedy: DDT. The 1950s saw uncontrolled spraying which checked the infestation in some places, but did not control the spread. In 1958, when DDT was banned, a less effective substitute Sevin was introduced and the government finally started a research programme to control the beasts.

Each year, the front of infestation moves westwards and southwards away from New England. Last year an estimated 13 million acres were defoliated, double the area in 1980. Estimates for the summer of 1982 are that yet another 13 million acres will be defoliated, that is, 50 per cent or more of the canopy will be destroyed by gypsy moths. Carried by vehicles, gypsy moths have been found in Ohio, North Carolina, South Carolina, Virginia, West Virginia, Wisconsin and Michigan and outbreaks have been reported in Salem, Oregon, and Santa Barbara, California. New England forests are not lumbered extensively any more, but as forests in the south and west become infested, the economic costs of the gypsy moth plague will grow.

The battle against the gypsy moth has been hindered by two things. One is the fact that no major industry is affected — thus in Maine, for example, the localized infestations are less important politically than the spruce budworm infestation, which has hurt that state's timber industry for some years. A second impediment is that the US environmental movement, with its power to express issues forcefully, and arouse public opinion, has been lukewarm to the gypsy moth issue. Many environmentalists are more concerned about the potential toxicity of Sevin, a spray commonly used against the gypsy moths, than they are about the forest damage *per se*. Moreover, the fact that lesser infestations do not destroy a forest completely and that most of the trees affected are only "ornamental", gives the environmentalists grounds to argue that the problem is not so urgent. One suspects that they may also hesitate to decry the gypsy moth because if the plague were indeed extensive is it not only a few logical steps to argue that we should return to the only known substance that combats it — DDT? Many environmentalists may not want to start down that slippery road.

So far, the gypsy moth story is one of the failure of science and technology. As the old saying goes, if we can go to the Moon, why can't we beat the gypsy moth? The chief microbial pathogen used so far, BT (*Bacillus thuringiensis*), is not nearly as effective as DDT: it must be sprayed over a large area requiring that an entire neighbourhood should get together and spray at once; if rain or other weather changes interrupt the spraying, it does no good. The other remedies tried so far have been of little effect, partly because, by now, the infestation is so severe. Research has recently focused on pheromones and other specific chemical lures, but the results are not yet broadly applicable.

The biology of the gypsy moth raises several interesting biological questions. The population dynamics of the gypsy moth are not understood, with rises of population followed by collapse, perhaps because of the weather. The physiology of the single American species may throw light on its susceptibility to parasites, perhaps the best hope for control. The relative scarcity of *Lymantria* species in Europe as compared with East Asia suggests that the genus originated in the Far East. The hope is that the work now under way will suggest a species-specific pesticide for a pest that the US Department of Agriculture acknowledges to be uncontrollable. The chemical company that first finds a cure will make a fortune. The search for parasites effective against gypsy moths has now turned to East Asia.

Upland oaks in New England heavily defoliated by the gypsy moth.

The pest also haunts Europe, but because the forests are less dominated by oak and are smaller and more precisely managed, it has not become the plague it is in the United States. Moreover, in Europe, the parasites that limit the gypsy moth population survive the winter whereas, when introduced into the United States, they die for lack of a suitable winter host.

China is the only part of the world that has not been thoroughly investigated for parasites that could solve the problem in the United States. It is to be hoped, therefore, that the Chinese government will be as cooperative as possible with the teams from the US Department of Agriculture which are there now, and plan to go again next year, to investigate gypsy moths in China, particularly in allowing parasites to be taken out of the country.

It is also to be hoped that the Reagan Administration will produce a better research policy on the gypsy moth than its parsimony so far suggests. The research programme begun by the Department of Agriculture in the 1970s may for practical purposes be abandoned as a consequence of the budget cuts. Aid to affected states on research and control, to which the federal government used to contribute 43 per cent, was cut for the present fiscal year to 12.5 per cent, although the Administration relented three weeks ago, and raised the proportion to 25 per cent. But in the coming financial year the Department of Agriculture will not be allowed to embark on any joint programmes with the states, although it may continue with some research of its own.

As New Englanders sit on their porches this spring, musing on how much the gypsy moth will devour of scrub oak or aspen trees on the West Coast, they can take some morbid pleasure in reminding themselves that their President thinks this is a local problem, and that his science adviser, for all his talk about the importance of science's interface with society, does not think the gypsy moth is of any priority for science.

Deborah Shapley

0028-0836/82/240531-01$01.00

© 1982 Macmillan Journals Ltd

CORRESPONDENCE

Herczynski's arrest

SIR — Many of your readers will have been dismayed to learn from Vera Rich's note (*Nature* 20 May, p.172) that Dr Ryszard Herczynski has been arrested in Warsaw for passing to two attaches from the US Embassy "materials damaging to the interests of the Polish state". Herczynski's numerous friends in this country know him as a vigorous and fearless critic of attempts by the authorities in Poland to inject political considerations into the conduct of scientific and academic affairs. That he has been critical of action taken under the present martial law, we can be sure; and it is probable that the materials alleged to be damaging to the interests of the Polish state were simply documents expressing such criticisms. Nevertheless, he could be sentenced to between three and five years' imprisonment (not internment) under martial law, with no appeal. This would be a terrible injustice.

Expressions of support for Herczynski from Western scientists may dissuade the Polish government from treating him harshly. It would be helpful to convey to those concerned that the essentially non-political causes defended by Herczynski — respect for the truth in scientific and academic matters and maintenance of the integrity of scientists — are also our causes, and that we do indeed care what happens to him. Quick action is desirable since Herczynski's trial may be held soon.

A letter of protest has been sent to the Polish authorities, signed by 25 scientists.

G.K. BATCHELOR

Department of Applied Mathematics and Theoretical Physics, Cambridge, UK

Japanese IQ

SIR — Reading Alun Anderson's comments on the study of IQ in Japan (*Nature* 20 May, p.180), the surprising find is not that it has improved with better nutrition, urbanization, improved social conditions and so on, but that under conditions comparable with those in the United States and Europe, it is 10–15 per cent higher. If this difference is genuine, I could offer an easily testable explanation.

It is well known that coaching can improve IQ results by 5–10 per cent. My conjecture is that every Japanese child undergoes a sort of coaching when (s)he learns to read and write his own language. Apparently it is so complex that only by the age of eleven can they be regarded as fully literate. Of course, the ideal IQ tests are language neutral, but my point is that the mastering of a difficult language plays a general IQ boosting role. This hypothesis is easily testable: one should compare the IQs of immigrant Japanese groups in the United States, say, who are literate in their language and matching groups who are not. Also, if the hypothesis is true, other members of the Chinese-related cultures should show higher IQ, under similar social and economic conditions.

Finally, competent linguists could perhaps find other languages of comparative complexity, the speakers of which could then be included in such inter-nation IQ comparison.

G. MAGYAR

Abingdon, Oxon, UK

Drugs and safety

SIR — I assume that the attack by M. Weatherall (*Nature* 1 April, p.387) on the toxicological testing of drugs was designed to provoke a response and was not intended to celebrate its publication date.

To propose the substitution of patient surveillance for toxicological testing is at first sight attractive because it places the emphasis firmly on men and women and not rats or bacteria. There is, moreover, undoubtedly a kernel of truth in some of his concern about unthinking and invalid testing, yet in selecting genetic toxicology to bear the brunt of his criticism he has chosen unwisely. His arguments show ignorance of current attitudes among thinking toxicologists and regulators. Moreover, the effects under consideration are, as I shall show, almost completely refractory to conventional surveillance.

Genetic toxicologists aim to detect potential mutagens and initiators of carcinogenesis. They would not expect that the result of a bacterial mutagenicity test alone could be extrapolated directly to man. Properly thought-out tests carried out early in the development of a drug can give an alert to possible problems ahead, problems that may be averted by chemical modification or that may require specialized tests to be undertaken when the drug is first given to humans. It would, of course, be verging on the criminal to give young patients with benign disease a drug known to be a potent mutagen in a wide variety of systems (including whole animal tests). Nevertheless, weak mutagens, or those active only in certain rather special situations, may well have a clinical future when all relevant circumstances have been considered. It is here (at the regulatory level) that the common sense advocated by Weatherall should be seen to operate.

To argue for surveillance and against testing for mutagenicity is particularly inappropriate since the deleterious effects of DNA damage are in general so long delayed that they would fail to be detected. Although sensitive techniques are now being developed that would enable the presence of DNA damage and other markers indicative of DNA damage to be detected in treated patients, their use is justifiable on economic grounds only where mutagenicity tests indicate possible risk.

Weatherall is right to try to educate the public and their regulatory bodies not to expect absolute safety from drugs any more than from surgical operations, but the surgeon introducing a new technique will have spent many hours experimenting on animals and will have a fair idea of possible risks before his first operation on a patient.

B.A. BRIDGES

MRC Cell Mutation Unit, University of Sussex, Brighton, UK

Back to basics

SIR — Many American scientists may be unconvinced that "Europe leads on (nucleotide) sequences" as suggested by your correspondent Robert Walgate in his article on the European Molecular Biology Laboratory (EMBL) computer library of nucleic acid sequences (*Nature* 15 April, p.596). American molecular biologists have for several years been successfully using several similar data banks including the Nucleic Acid Sequence Data Base organized by Margaret Dayhoff's group at Georgetown University Medical Center. This sequence library, in addition to containing several programs for sequence analysis, currently includes 746,000 nucleotides (compared with the 600,000 reported in your article for the EMBL library).

Apart from the issue of nucleotide quantity, the sample entry provided in your article, while perhaps not typical, does not reflect favourably on the EMBL library as regards quality. The entry gives no indication that the MOPC41 kappa gene sequence presented was not determined by the authors of the entry reference; the sequence was in fact copied by them in their paper to compare with a sequence of another gene which they had determined. In transcribing the MOPC41 sequence these authors erroneously inserted a nucleotide (position 189) which now appears in the EMBL sequence. This trivial error illustrates the importance of relying on original sources of sequence data, a policy of the Georgetown data base. The correct sequence is included in the corresponding entry from the Georgetown data base (see below) along with the correct primary reference (Seidman *et al. Nature* **280**, 370; 1979).

JONATHAN SEIDMAN
EDWARD E. MAX

Department of Health & Human Services, National Insitutes of Health, Bethesda, Maryland, USA

```
KVMS41
Ig kappa chain V region germline gene VK41 - Mouse

Seidman, J.G., Max, E.E., and Leder, P., Nature 280, 370-375,
    1979 (Residues 1-664)
     Residues            Feature
     120-174,            Protein: Ig kappa chain precursor V
     303-598             region MOPC 41
     175-302             Intron

Composition:   168 A,   153 C,   139 G,   204 T
Length:        664

                10         20         30         40         50         60
  1 CGTGACCAAT CCTAACTGCT TCTTAATAAT TTGCATACCC TCACTGCATC GCCTTGGGGA
 61 CTTCTTTATA TAACAGTCAA ACATATCCTG TGCCATTGTC ATTGCAGTCA GGACTCAGCA
121 TGGACATGAG GGCTCCTGCA CAGATTTTTG GCTTCTTGTT GCTCTTGTTT CAAGGTTAAA
181 ATGAAACTAA AATTGGGAAT TTTCCACTGT TTCCAACTGT GGTTAGTGTT GACTGGCATT
241 TGGGGGATGT CCTCTTTTAT CATGCTTATC TATGTGGATA TTCATTATGT CTCCACTCCT

               310        320        330        340        350        360
301 AGGTACCAGA TGTGACATCC AGATGACCCA GTCTCCATCC TCCTTATCTG CCTCTCTGGG
361 AGAAAGAGTC AGTCTCACTT GTCGGGCAAG TCAGGACATT GGTAGTAGCT TAAACTGGCT
421 TCAGCAGGAA CCAGATGGAA CTATTAAACG CCTGATCTAC GCCACATCCA GTTTAGATTC
481 TGGTGTGCCC AAAAGGTTCA GTGGCAGTAG GTCTGGGTCA GATTATTCTC TCACCATCAG
541 CAGCCTTGAG TCTGAAGATT TTGTAGACTA TTACTGTCTA CAATATGCTA GTTCTCCTCC

               610        620        630        640        650        660
601 CACAGTGATA CAAATCATAA CATAAACCCC ATGAAAGTAG AAATGAGAGG CTGGGCTGCT
661 CTGA
```

© 1982 Nature Publishing Group

NEWS AND VIEWS

Female mate choice and runaway sexual selection

from Paul H. Harvey and Stevan J. Arnold

IN *The Descent of Man and Selection in Relation to Sex*[1], Darwin explained how various traits usually found in only one sex — such as the peacock's tail or the deer's antlers — might have evolved. He thought that such traits resulted from one of two selective processes that favour access to mates: either selection for weapons used in combat between males, or selection for adornments used to attract females (Figs 1, 2). Darwin observed that the adornments are most clearly seen in males of species, for example, birds of paradise, where females choose from a group of displaying males. But Darwin was aware of a difficulty in his theory that has never been fully resolved: the characters favoured by females often appear to hinder the survival of males. Why should females choose to mate with those males least likely to be favoured by natural selection? Recent analytical work formalizes a model suggested by R.A. Fisher and shows how selection can favour female mating preferences for characters that actually lower male viability, even to the point of driving a population to extinction.

Fisher[2] described a simple model for female choice which had a surprising outcome — male characters which impair survival could be selected for and maintained in a population simply as a consequence of female choice. He considered species with heritable variation for both a male character, say tail size, and female preference for choosing mates on the basis of tail size. If environmental change led natural selection to favour an increase (or decrease) in male tail size, then, over the generations, male tail size would change. However, so would female preferences. The important idea here is that males with longer tails would tend to be mated by females who prefer longer-tailed males. The male offspring of those females will not only inherit their fathers' long tails and be favoured by natural selection but will also inherit their mothers' preference genes and pass them on to their daughters, thus causing an increase in the frequency of those preference genes. This genetic covariance or linkage disequilibrium is a result (not an assumption) of Fisher's model. Modification of male tail size will, therefore, proceed under the influences of both natural selection and sexual selection.

Ultimately a balance between the forces of natural and sexual selection might be expected to lead to an equilibrium male tail size, but Fisher[2] argued that "as long as there is a net advantage in favour of further plumage development, there will also be a net advantage in favour of giving to it a more decided preference". Fisher termed this process 'runaway selection' and went on to assert that in the absence of "severe counterselection . . . *it is easy to see* that the speed of development will . . . . increase with time exponentially, or in geometric progression" (our italics). He suggested that the process might only be checked when males had tails so long that they would be extremely unlikely to survive and females preferring to mate with them would be left without a mate.

Fisher never published a formal population genetic model of the runaway process and, since his writing style was never lucid, later workers have often chosen to add assumptions to the model. One common variant[3,4] claims that the runaway process is a consequence of females choosing males on relative rather than absolute criteria — for example, from one extreme of the distribution of observed male tail sizes in the contemporary population rather than males with particular sized tails.

There has been a clear need for explicitly formulated models that would describe the joint evolution of male characters and female choice and that would, at the very least, specify some conditions under which the runaway process does or does not operate. In a series of papers published in the 1960s and 1970s and summarized in a book[5] published in 1980, O'Donald examined several models which incorporated genetic loci coding for male characters and female choice. These two-locus diploid systems cannot be treated analytically, but computer simulations show that the runaway process would occur. However, Fisher "describes the

Fig. 1 (right) Strepsiceros Kudu. One category of sexually selected characters confers a direct advantage in male–male combat over access to mates. Darwin was aware of numerous examples and had no difficulty in explaining them. Male weaponry is often more markedly developed among polygynous species. **Fig. 2** (below) A second category of sexually selected characters is particularly evident in males of lek species, such as the bird of paradise, where females choose from among a group of displaying males. The males play no part in defending the nest or feeding the young. But why do females appear to choose to mate with males whose plumage must be disadvantageous in terms of natural selection?

Paul H. Harvey is a lecturer in Biology at the University of Sussex Falmer, Brighton BN1 9RH and Stevan J. Arnold is an Associate Professor of Biology at the University of Chicago.

0028-0836/82/240533-02$01.00

© 1982 Macmillan Journals Ltd

evolution of mating systems in terms Darwin himself might have used"[5] and not all his claims are validated in O'Donald's population genetic treatment of his model — the runaway process does not increase at a geometric rate.

Working with a haploid two-locus analytical model and diploid two-locus simulations, the recent treatments by Kirkpatrick[6] also demonstrate the runaway process. But it is evident from Kirkpatrick's treatment that both Fisher and O'Donald were wrong to conclude that "preferences start to evolve only if females prefer those male phenotypes that are advantageous in natural selection"[5]. Indeed, female preference genes that favour more viable males do not necessarily spread through the population, and neither do the more viable males! Why should this be so? The key to the answer lies with Kirkpatrick's finding that when the male trait gene frequency is plotted against female choice gene frequency then there is a line of genetic equilibria: for every frequency of preference there is a frequency of the male trait at which the system is in equilibrium. There is not a single equilibrium point because any viability loss experienced by long-tailed males can, with some female preference, be exactly balanced by a mating advantage. When females preferring long-tailed males are common in the population then sexual selection will maintain a high frequency of the long-tailed males. When the preference gene is not at fixation, however, a low frequency of shorter-tailed males may be maintained in the population to service those females who prefer mating with them. The process is clearly frequency dependent.

What this means is that when the same male fitness and female choice functions are used while initial trait and preference frequences are varied then different points on the line of equilibria are reached. Once populations reach the line there is a prospect for additional indeterminate evolution — genetic drift can carry small populations along the line of equilibria. Finally, it is clear from the models that females need not choose males on relative rather than absolute criteria (see above and ref. 7) for the male's tail to evolve to a maladaptive size.

Kirkpatrick's analysis independently followed a treatment of the problem by Lande[7] which, more realistically, assumed that both male traits and female mate preferences are under polygenic control. Lande modelled such a system with male traits and female preferences treated as normally distributed variables. Although we might expect selection to deplete heritable variance in both traits, Lande points out that this is not necessarily so[8]; at some point depletion by selection will be balanced by enrichment of variation by mutation and recombination. If we assume that variances in male trait and female preference metrics remain the same

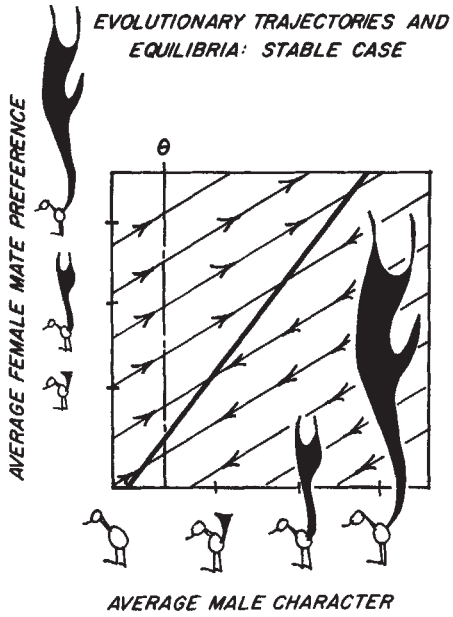
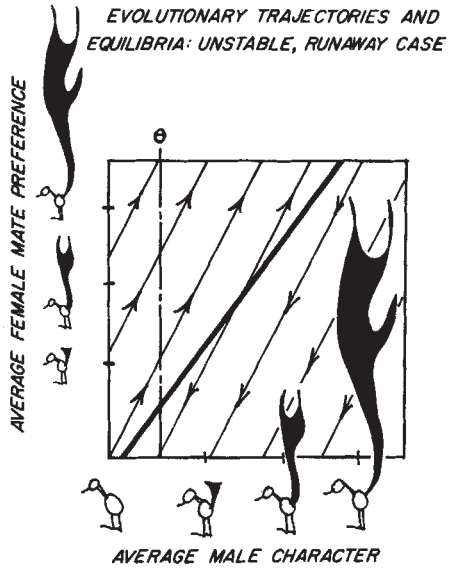


Fig. 3 Females choose mates on the basis of tail size. The light lines depict possible trajectories for the joint evolution of male tail size and female preferences, and the arrows indicate the direction taken. (The male tail size that is optimal for viability is indicated by the symbol θ.) The heavier line connects the equilibrium states at which the males tail length and the females mating preference cease to undergo evolutionary change (other than by random drift along the line). Whether the evolutionary process leads to stable or unstable equilibria depends on genetic parameters and the strength of viability selection.



between generations, but that character averages change as consequences of natural and sexual selection, then Lande's models also reveal a line of equilibria. For any st: ength of female preference there is a male trait size such that the system is in equilibrium. Lande demonstrates the line of equilibria in three models assuming quite different rules of female choice. Again, there is no direct selection on female mating preferences, which evolve as a genetically correlated response to selection on males.

Lande's formulation allows a different interpretation of what Fisher meant by runaway selection. With the two-locus models, populations perturbed slightly from the line of equilibria always return there (although not necessarily to the same point) and the runaway process ends. However, in Lande's models, evolutionary trajectories may lead away from the line of equilibria so that both male trait and female preference can evolve "away from the line at a geometrically increasing rate" (Fig. 3), as Fisher originally claimed. The qualitative conditions for the line of equilibria to be unstable are that females show much heritable variation in mate preference and that there is weak natural selection on the male character. Lande points out that his models help to explain why males of closely related species may differ in a "substantially nonadaptive and random pattern" while the females are phenotypically similar, and also that "random genetic drift may be an important factor promoting speciation by sexual isolation and the evolution of sexual dimorphism".

The similar conclusions reached by models with such dissimilar genetic assumptions underscores the generality of the conclusions. Lande's model, in particular, is general in terms of its assumptions yet clear predictions follow from it[9]. One other idea based on females choosing bizarre males precisely because the male traits constitute handicaps to survival[10] has gained questionable credence among some ethologists who seem not to understand that population geneticists have been unable to model the process successfully. O'Donald[5] presents a detailed analysis of the topic.

The model first sketched by Fisher 67 years ago[11] is only now becoming understood and its widespread implications appreciated. There is no reason why the runaway process should be restricted to plumage characters — vocalizations or pheromone production are other possible candidates. Shortly we might expect relevant parameters to be measured and the efficacy of the theory to be tested. In any event, the finding that, under a variety of circumstances, sexual selection can result in a decline of male viability and thus contribute to population extinction is an important one for evolutionary biology. □

1. Darwin, C. *The Descent of Man and Selection in Relation to Sex* (Murray, London, 1871).
2. Fisher, R.A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
3. Trivers, R.L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B.) (Aldine, Chicago, 1972).
4. Halliday, T. in *Behavioural Ecology, an Evolutionary Approach* (eds Krebs, J.R. & Davies, N.B.) (Blackwell, Oxford, 1978).
5. O'Donald, P. *Genetic Models of Sexual Selection* (Cambridge University Press, 1980).
6. Kirkpatrick, M. *Evolution* **36**, 1 (1982).
7. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3721 (1981).
8. Lande, R. *Genet. Res. Camb.* **26**, 221 (1976).
9. Arnold, S.J. in *Mate Choice* (ed. Bateson, P.P.G.) (Cambridge University Press, 1982).
10. Zahavi, A. *J. theor. Biol.* **53**, 205 (1975).
11. Fisher, R.A. *Eugenics Rev.* **7**, 184 (1915).

© 1982 Nature Publishing Group

'Style' in prehistoric artefacts

from Paul Mellars

VARIATIONS in the forms of hunting weapons have always held a peculiar fascination for archaeologists. In North America, for example, the whole taxonomy of the 'Palaeoindian' cultures depends heavily on the elaborate patterning in the various forms of stemmed and fluted projectile points, whilst in Europe the same preoccupation can be seen in the classification of a wide range of industries from the Upper Palaeolithic period to the Bronze Age.

The tacit assumption underlying these studies has always been that the patterning represented in the projectile point forms reflected in some way the social groupings of the human communities themselves — that the archaeological 'cultures' bore at least some resemblance to the kind of ethnic or cultural groupings which a modern ethnographer or anthropologist might recognize. From the archaeological evidence alone, of course, any correlations of this kind must remain essentially a matter of faith, and it is only with the recent surge of interest in 'ethno-archaeological' studies that the prospect of testing these correlations has emerged.

By far the most important study to date was reported by Polly Wiessner, of the Max Planck Institut für Humanethnologie, at the recent annual conference* of the Prehistoric Society. For several years Wiessner worked among a number of bushman groups (collectively known as the 'San') who still live predominantly by hunting and gathering in the northern and central parts of the Kalahari Desert. Her main aim was to investigate 'style' as expressed in the material culture of these groups, focusing particularly on how far stylistic variation in the different types of artefacts could be correlated with the various linguistic divisions and other forms of social groupings of the present bushman communities. The analysis was conducted at several levels, ranging from variation between individual craftsmen, through the successively larger 'band' and 'band-cluster' groupings, to the level of the three major linguistic divisions (the !kung, G/wi and !xo) which represent the most discrete and ancient social divisions within the societies studied.

From an archaeological perspective, the most instructive results were obtained from a detailed study of the metal-tipped hunting arrows, which are not only among the most highly valued possessions of the bushmen, but are also still used to secure much of the daily food supply. Most important was the discovery that whilst the forms of these arrows showed a good deal of variation from one hunter to another (and sometimes between different batches of arrows manufactured by the same hunter), it was only at the level of the major linguistic grouping that any consistent and statistically significant patterns of variation could be observed. Curiously, variation at this level was apparent not only in relatively minor details of the shapes of the arrow points, (for example, the convexity of the sides or the precise form of the barbs) but also in features that might be expected to have a significant effect on the function of the arrows. Thus, the arrowheads manufactured by the !kung were consistently smaller than those made by the other two groups, whilst those produced by the !xo had distinctively broader, flatter tips. In these features there was no detectable overlap in the forms of the arrows manufactured by the three groups, so individual arrows could easily be visually assigned to one group or another. Despite these variations, the different forms of arrows were used to kill almost exactly the same species of game in all the societies examined.

Most archaeologists would probably argue that these results conform exactly to expectations — that the element of style in the form of the arrow points was serving to differentiate clearly separate social groupings which would have meaning not only to an anthropologist but also to the bushman groups themselves. Wiessner clearly regards this as one of her most important findings, and has referred to this form of patterning as 'emblemic style', that is, the use of stylistic variation as a symbol or 'emblem' which serves to link the individual with some specific, clearly defined social group.

One of the most intriguing implications of her study, however, is that in certain other spheres of material culture, style can operate in a totally different way. The clearest illustration of this was provided by her analysis of the elaborate beadwork used by bushman women to decorate clothing and other personal possessions. In contrast to the clear patterning observed in the arrow-point forms, the patterns of beadwork showed surprisingly little clear-cut variation from one bushman group to another, and instead displayed a bewildering array of variation in the products of the individual women within each group. In other words, style in this case was focused much more at the personal level, than at the level of the local or regional group. Wiessner has therefore referred to this form of highly individualistic style as 'expressive style'.

Wiessner suggests that much of the contrast between the two forms of style can be explained in terms of the particular functions which the different categories of artefacts served in bushman life. The items decorated with beadwork were largely personal possessions which rarely changed hands between individuals and which appeared to play little part in the social relationships maintained between different bushman groups. Arrows, on the other hand, had a much more dynamic role in social life. Not only were arrows frequently exchanged between individual hunters, but the forms of the arrows could potentially be used as a guide to the ownership of killed or wounded game. As Wiessner points out, the ownership of killed game has virtually no significance at the level of the individual hunter, since the strongly egalitarian ethic of bushman society demands that meat from any animal should be widely distributed between members of the local residential group. Even at the level of the separate bands or band-clusters, the question of ownership is of limited relevance, since most of the bands within each major linguistic group maintain a close alliance in economic matters which allows relatively free access to the territories (and therefore food resources) of other groups, and which frequently leads to the movement of individual hunters from one band to another. At these levels of organization, therefore, there is little need for any clear identification of ownership of killed animals, and indeed the uniformity in the form of the arrows could be viewed as an important way of symbolizing and reinforcing the economic and social ties between the different groups. At the level of the linguistic groups, however, social divisions are much more sharply defined. Social or economic exchanges taking place across these boundaries are generally more formalized, and more dependent on a clear recognition of the rights and affiliations of the groups or individuals involved. Wiessner suggests that the clear differentiation in the arrow forms at this level could be important not only economically (by clarifying the ownership of wounded animals) but also in identifying the sources of any aggression or personal conflict across group boundaries. The clear messages encoded in the forms of the arrows thus become an important element in the patterns of 'boundary maintenance' between the societies involved.

Clearly, the archaeological implications of this study need to be treated with caution. Bushman societies have undergone radical transformations in recent years and it is inevitable that many aspects of both bushman life and bushman technology are still adjusting to these changes. Nevertheless, there are several lessons implicit in the bushman data to which the archaeologist should take heed. As a number of authors have recently emphasized (for example, Wobst in *Univ. Michigan Mus. Anthrop., Anthrop. Pap.* **61**, 1977; and Hodder in *The Spatial Organization of Culture*, Duckworth, 1978), 'style' is not a simple univariate

*A conference on 'The Archaeology of Hunter-Gatherers' was held in London on 20–21 March 1982.

P.A. Mellars is in the Department of Archaeology, University of Cambridge, Downing Street, Cambridge CB2 3DZ.

0028-0836/82/240535-02$01.00

© 1982 Macmillan Journals Ltd

phenomenon which can be equated in a simple, one-to-one way with the extent of 'interaction' between separate groups or individuals; it can clearly operate very differently in different social contexts, and can lead to a variety of contrasting and interacting patterns in different spheres of material culture. Variations in the forms of hunting weapons may after all provide one of the most sensitive indicators of group affiliation among prehistoric hunting and gathering communities, but to assume that this would necessarily be true in all social contexts would be stretching the evidence too far. Only by clearly understanding the social mechanisms by which these variations in material culture come into being can we reliably interpret the data from the surviving archaeological record. It is in this context, of course, that the ethnoarchaeological studies carried out by Wiessner and others are of central importance to the understanding of human behaviour in the past. □

Total eclipse of June 1983

from A.D. Fiala

NEXT June 11, nature will provide one of her most spectacular sights, a total eclipse of the Sun. The eclipse will be of long duration and visible from accessible land sites where there is every prospect of good weather — a favourable combination that has not occurred since June 1973 and will not occur again until July 1991. The path of totality will cross land mostly in Indonesia, but also in Papua New Guinea and a few small islands. Maximum duration of totality will reach over 5 minutes 15 seconds, at a point midway between the islands of Java and Sulawesi. However, on both islands duration in excess of five minutes will occur, with the Sun well up in the sky.

The path of totality will start in the south Indian Ocean at sunrise. For the first hour the umbra will sweep over water, then fall upon Christmas Island. It will then travel the Indonesian archipelago, crossing the centre of Java, the south of Sulawesi and the south of New Guinea, with Ujung Pandang and Port Moresby lying nearly on the central line (see the map). On Java, several large cities lie in the path, including Jogjakarta, Surabaya and Semarang. Bosscha Observatory, the closest to the equator of the world's major observatories, will lie outside the path of totality, however, and will see only 97 per cent partial phase.

The weather prospects for the eclipse are excellent, according to Canadian meteorologist Jay Anderson (*Newsletter of the Bull. R. astr. Soc. Can.*, June 1981). Monsoon conditions give a probability of clear sky of over 65 per cent on the north-east coast of Java and south-west coast of Sulawesi, and nearly 90 per cent at Port Moresby. The sunspot cycle will be midway between maximum and minimum, so intermediate coronal structure and activity may be expected. The eclipse will last long enough for the sky to darken, making it easier to see faint features, particularly from high altitude.

Many astronomers and physicists are laying provisional plans as, indeed, is the Indonesian Government, which is hoping to attract tourists to both the eclipse and the hundredth anniversary of the eruption of Krakatoa that falls at the same time. Experiments are being planned to measure the size of the Sun, to probe the composition and structure of the solar atmosphere and to look for solar effects on the Earth's atmosphere and on the behaviour of living organisms.

A collaborative team from the US Naval Observatory, the Computer Science Corporation, NASA Goddard Space Flight Center and the Astronomical Society of Canberra plans to man at least one station near each edge of the path, in order to measure the solar diameter. This procedure has been carried out at many eclipses, since the original measurements in 1715 by Halley. So far, the results indicate that there is a variation in the size of the Sun, but too few measurements have been gathered to distinguish between monotonic and pulsating variation. The long duration of this eclipse should give the opportunity for observing an extraordinarily large number of small limb features. The observations are not at all difficult, requiring only careful timing and concentration.

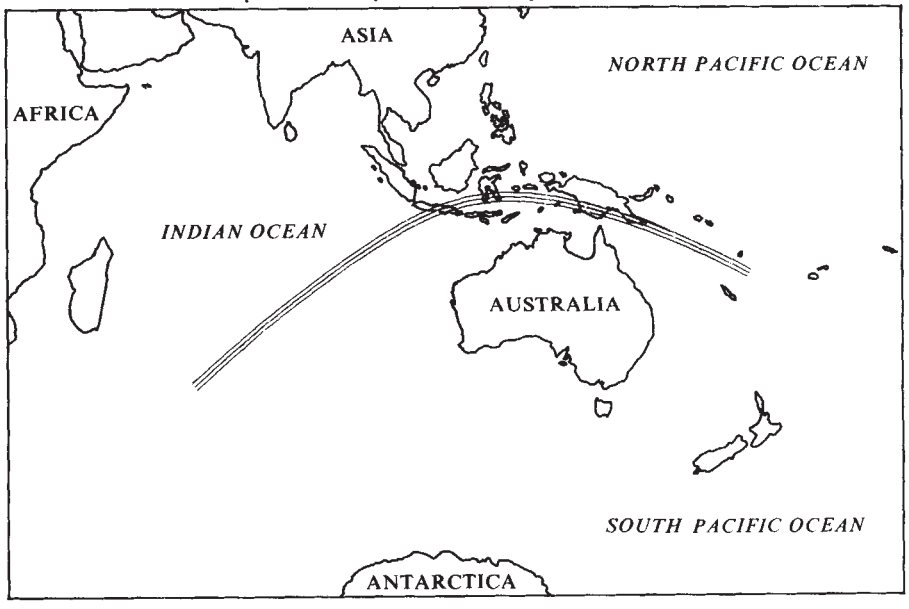
Nowadays at any major eclipse there are several airborne observers. NASA's Kuiper Airborne Observatory may be flying along the track over the Indian Ocean, using the Moon as a high-resolution image dissector along the limb to study the limb brightness variation in sub-millimetre wavelengths.

On the ground, several US experimenters hope to be present. A team led by D. Landman, Institute of Astronomy, Hawaii, hopes to continue spectrophotometry of the weak emission lines and continua of ordinary prominences to provide tests of theoretical models. Investigative teams from Kitt Peak, Sacramento Peak and Williams College would like to study in different ways the motion of material in the inner corona, while a group from Iowa State University would study the motion of dust in the outer corona.

Turning the Sun 'off' for several minutes should produce some significant effects in the Earth's atmosphere in the wake of the eclipse path. At least four teams from the US are hoping to study these effects. Two of them would attempt to refine our knowledge of the atmospheric turbulence causing shadow bands and another would look for effects of the wake as far away as India. Observing teams are also expected from Japan and Australia, and quite likely from India, Europe and Canada as well. Indonesia has created a national committee on the eclipse.

This eclipse is the first of three total eclipses to occur in the region over a five-year period. Interestingly, the second is also the next total eclipse, on 22–23 November 1984, and its path intersects with the 1983 eclipse in the Gulf of Papua. However, no experiment has yet been proposed which would make use of this circumstance. The third eclipse will be on 18 March 1988 over Sumatra and Borneo.

The path of totality of the total eclipse of June 1983.



A.D. Fiala is at the US Naval Observatory, 34th and Massachusetts Avenue, NW Washington, DC 20390.

0028-0836/82/240536-01$01.00

© 1982 Macmillan Journals Ltd

Ionospheric hole caused by rocket engine

from Michael J. Rycroft

AMONG the most interesting of the wide variety of active experiments[1] that can be carried out to study the tenuous upper atmosphere and ionosphere are those that are 'accidental'. One such experiment, perhaps better termed an 'experiment of opportunity', is the operation of a rocket engine at heights above 200 km. The Saturn V launching of Skylab on 14 May 1973, which halved the total electron content of the ionosphere for three hours[2,3], created considerable interest in the effects of rocket exhaust gases on the ionosphere. Further measurements of an artificially induced 'ionospheric hole' were made during the Atlas-Centaur launch on 20 September 1979 of HEAO-C[3,4]. Michael Mendillo and Jeffrey Baumgardner, of Boston University, report[5] the first results of the latest experiment of this type, that for the Atlas-F launch of the NOAA-C weather satellite on 23 June 1981. Launched from Vandenburg Air Force Base, California, at around 03.00 local time, the rocket engine burned to an altitude of 434 km, well into the topside F-region of the ionosphere.

Because the rocket exhaust gases contain highly reactive molecules such as H_2, H_2O and CO_2, reactions with atomic oxygen ions (which are dominant in the F-region) produce molecular ions such as H_2O^+ and H_3O^+. The rates of the reactions are between a hundred and a thousand times faster than those which normally occur with ambient atmospheric nitrogen and oxygen molecules. Thus the atomic ion plasma is rapidly converted to a molecular ion plasma which then recombines quickly with ionospheric electrons, removing ionization and creating the local ionospheric hole. As the rocket proceeds and the plume of exhaust gases diffuses through the atmosphere, the electron content of a column through the F-region, both bottomside and topside, is depleted. The dissociative recombination reactions of the molecular ion plasma, which create the hole, produce oxygen atoms in the singlet D excited state. As these fall into triplet P states, red airglow is emitted, primarily at a wavelength of 630 nm. The position where the airglow is emitted is determined by the time scale of the diffusion of rocket exhaust species through the atmosphere.

The two important quantities to be observed in such experiments are the airglow and the total electron content (TEC) of the ionosphere. In the June 1981 experiment, the TEC was found by measuring the Faraday rotation of the plane of polarization of 136 MHz radio beacon signals from the ATS-1 geostationary satellite over the Pacific Ocean, the groundsite for such observations being chosen such that the ray path passed through the rocket's exhaust plume at 350 km altitude. Four minutes after launch, the TEC decreased rapidly, and reached a steady value some three minutes later; the TEC was found to decrease by 1.7×10^{17} electrons per m^2. From the same site narrow-band optical observations were made using a low-light level image-intensified photographic system with a 60° field of view. As the TEC decreased, an expanding shell of airglow with an intensity of several kiloRayleighs was observed; in the horizontal plane, this appeared rather like a 'smoke ring' with a radius of up to about 1,200 km. As the shell expanded more, its intensity decreased on a time scale of several minutes.

What can be deduced from such observations of the expanding shell of rocket exhaust molecules? First, the diffusion coefficient can be estimated to be $3 \times 10^7 m^2 s^{-1}$. This is consistent with the lightest species of exhaust gases (H_2) diffusing through an ambient atomic oxygen atmosphere, with a temperature of 1,000K at 350 km altitude. Second, knowing that the Atlas-F rocket ejects 10^{27} H_2O molecules per s (about 50 kg s^{-1}), the observation that the airglow emission was maximal two minutes after the initiation of the hole can be interpreted. It is found that 14 per cent of the plasma recombinations yield oxygen atoms excited in the singlet D state. Also, it is shown that, to first order, quenching of the excited oxygen atoms by molecular species is not important.

Further analysis of this and similar experiments can be expected in the future, and other chemical release experiments are planned, for example from the Space Shuttle. Even the operation of the Shuttle's engines or the dumping of water produced by its fuel cells could lead to interesting results. Such active space plasma experiments conducted in the ionosphere are not only of scientific interest, but also have practical significance for various aspects of radio communications. □

Michael J. Rycroft is Head of the Atmospheric Sciences Division of the British Antarctic Survey, NERC, Madingley Road, Cambridge CB3 0ET.

1. Rycroft, M.J. *Nature* **287**, 7 (1980).
2. Mendillo, M., Hawkins, G.S. & Klobuchar, J.A. *J. geophys. Res.* **80**, 2217 (1975).
3. Mendillo, M. *Adv. Space Res.* **1**, No.2, 275 (1981).
4. Mendillo, M., Rote, D. & Bernhardt, P. *EOS Trans. Am. geophys. Un.* **61**, 529 (1980).
5. Mendillo, M. & Baumgardner, J. *Geophys. Res. Lett.* **9**, 215 (1982).

Coastal plants take to the road

from Peter D. Moore

IT has long been recognized that the spreading of salt on roads in Britain during the winter is having a progressive influence on the vegetation of our roadside verges[1]. A similar problem has been described from the United States back in 1969, when it was estimated that 6×10^6 tons of salt were applied annually to roads in the northern states[2]. Up to 10^6 tons are used annually in Britain[3], and locally this often means that salt is being added to the soil of these verges at a rate of 3–4 kg m^{-2} yr^{-1} (in the United States, values of up to 6 kg m^{-2} yr^{-1} have been recorded[2]).

If we compare these figures with those obtained from studies of natural salt spray on coastal vegetation, such as maritime cliff grasslands, then the ecological impact of the salt input can be appreciated more fully. In his studies of the Lizard peninsula, Cornwall, Malloch[4] determined the annual 'salt' (Na, K, Mg, Ca, Cl) deposition in rain gauges to be approximately 0.07 kg m^{-2} yr^{-1}. In only one location did he record a value as high as 0.86 kg m^{-2} yr^{-1} of sodium, which is approximately equivalent to 2.6 kg m^{-2} yr^{-1} of total salt. The results are in keeping with those from sites in western Ireland 15 km from the sea[5] (0.04 kg m^{-2} yr^{-1}), in Lancashire 2 km from the sea[6] (0.01 kg m^{-2} yr^{-1}) and in Cheshire 40 km from the sea[6] (0.004 kg m^{-2} yr^{-1}). It is evident, therefore, that our roadside verges are receiving a considerably greater (often by a factor of 50) input of salt than our most exposed cliff grasslands, and it is not surprising that the effects of such treatment are being increasingly felt.

The application of salt to soil has a strong influence on its structure, and leads to clay flocculation and consequent loss of aeration. It also, naturally, raises the osmotic potential of the soil water, making it difficult for plants which lack physiological adaptation to absorb water. There are two main consequences for the vegetation: first, there will be a strong selective pressure against those individuals within populations that lack the requisite adaptations; and second, halophytic, maritime plant species will be able to invade these inland habitats.

Relatively little data are available on the effects of salt application upon the general

Peter Moore is in the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF.

0028-0836/82/240537-02$01.00

© 1982 Macmillan Journals Ltd

balance of grassland species but it seems that white clover (*Trifolium repens*) and meadow grass (*Poa pratensis*) may suffer more than red fescue (*Festuca rubra*), rye grass (*Lolium perene*) and crested dogstail grass (*Cynosurus cristatus*)[7]. Selection of specifically adapted genotypes within populations is difficult to demonstrate in the field, but experimental studies suggest that this will be an active process[8].

The second consequence — the invasion of typically coastal species of plant — is more obvious to the observer and hence has been studied in more detail. In 1976, Mathews and Davison[9] surveyed the Newcastle-upon-Tyne area and found several such species (such as *Aster tripolium, Puccinellia distans, Plantago maritima* and *Sueda maritima*) on roadside verges as much as 13 km from the sea. The salt marsh grass (*Puccinellia distans*) has now been recorded on verges extensively throughout northeastern England, in the Midlands, East Anglia and Kent[10]. Other species have not kept up with the rapid spread of *Puccinellia,* though the sea plantain (*Plantago maritima*) and the sea spurrey (*Spergularia marina*) are still advancing gradually. The sea aster (*Aster tripolium*), however, appears to have failed in its bid for a place in the new saline grasslands; seed set has been poor and populations are contracting rather than expanding. Seed dispersal techniques could be the key to success in this invasion. *Puccinellia, Plantago* and *Spergularia* are low-growing plants and their seeds could easily be carried in the mud adhering to vehicle tyres, while *Aster* seeds are produced well above the ground and rely on wind dispersal.

A survey of roadside vegetation in Norfolk has also recently been conducted by Bull[11]. Permanent recording plots were set up five years ago and the species present recorded annually since then. Both extinctions and immigrations have been observed during this time and Bull regards many of the new immigrants (for example, *Polygonum aviculare, Artemisia vulgaris* and *Rumex obtusifolius*) as salt tolerant. The difficulty in interpreting this kind of data is the number of variables involved (frequency and time of mowing in particular); controlled experimental studies are evidently needed.

The invasion of our roadside verges by halophytes is thus proceeding in much the way that the 'railway flora' developed during Victorian times[12]. One difference is that the nature of the current change will be less conspicuous. Extinctions may be more abundant than immigrations, but these are less easily noticed by casual observers, and genetic modifications within our roadside plants may also be overlooked. We may be sure, however, that there are subtle changes afoot in our 1,500 km^2 of roadside verges[13]. □

1. Moore, P.D. *Nature* **263**, 189 (1976).
2. Westing, A.H. *Phytopathology* **59**, 1174 (1969).
3. Davison, A.W. *J. appl. Ecol.* **8**, 555 (1971).
4. Malloch, A.J.C. *J. Ecol.* **60**, 103 (1972).
5. Sparling, J.H. *J. Ecol.* **55**, 1 (1967).
6. Allen, S.E., Carlisle, A., White, E.J. & Evans, C.C. *J. Ecol.* **56**, 479 (1968).
7. Bannister, P. *Introduction to Physiological Plant Ecology*, 198 (Blackwell Scientific, Oxford, 1976).
8. Rush, D.W. & Epstein, E. *Pl. Physiol.* **57**, 162 (1976).
9. Mathews, P. & Davison, A.W. *Watsonia* **11**, 146 (1976).
10. Scott, N.E. & Davison, A.W. *Watsonia* **14**, 41 (1982).
11. Bull, A. *Trans. Norfolk Norwich Nat. Soc.* **26**, 23 (1982).
12. Lousley, J.E. in *The Flora of a Changing Britain* (ed. Perring, F.) 73 (Classey, Hampton, 1970).
13. *Amenity Grasslands, the Need for Research* (NERC report, 1977).

Seawater under extreme pressure

from Michael Whitfield

UNTIL recently, the experimental difficulties involved in performing precise measurements at high pressures have ensured that relatively little information was available about the properties of seawater and the reactions of its constituents in ocean conditions. The available information has been scattered throughout the literature, uncorrelated and accessible only to the *cogniscenti.* Two recent publications[1,2] indicate that this picture is changing rapidly.

Most experiments are carried out at atmospheric temperature and pressure, but it is a harsh reality that mean ocean conditions lie closer to 4°C and 200 atm, and more extreme conditions are not uncommon. The pressure generated by the overburden of water has a significant influence on the density of seawater and hence on the stability of the water column and the direction and rate of water movement. Furthermore, minerals such as calcium carbonate and silica, generated in surface waters by biological processes, become progressively more soluble as they settle through the ocean, encountering waters at increasing pressures. Consequently the recycling of these minerals and the many trace elements associated with them is strongly dependent on the pressure sensitivity of the dissolution reactions. The solubility of calcium carbonate in the deep oceans is of particular concern since the dissolution of this mineral is likely to provide the final sink for the excess carbon dioxide currently being injected into the atmosphere by the combustion of fossil fuels. Another reason for growing interest in these measurements is the discovery of rich and diverse communities of organisms surrounding warm hydrothermal oases along the mid-oceanic ridges which are subjected to pressures of several hundred atmospheres.

To appreciate the significance of the recent developments one must look first at the problem of defining the equation of state for seawater, which enables the density of seawater to be calculated at a particular temperature, pressure and salinity. These parameters can be accurately measured and the calculated density fields then used to elucidate oceanic circulation patterns.

To establish the new equation of state for seawater, techniques were developed for the precise measurement of solution density over a wide range of temperature and pressure. Density measurements at atmospheric pressure for solutions of individual electrolytes found in seawater can be used to estimate the contributions made by the component ions to the solution volume (the ionic molal volumes). Similar measurements made at high pressure provide estimates of the ionic molal compressibility. The influence of pressure on a chemical reaction can be estimated from the accompanying volume change — a reaction which produces a decrease in volume being favoured by the application of pressure.

Millero[2] has taken advantage of recent density measurements at atmospheric pressure and at high pressure to compile data on ionic molal volumes and compressibilities over the temperature and pressure range of interest to oceanographers (0–50°C, 1–1,000 atm). Using this information he has calculated the effect of pressure on the solubility of a range of mineral phases (including silica, strontium carbonate and calcium carbonate) and shown that the calculated values are in good agreement with available experimental data. Furthermore, by taking into account the molal compressibilities he was able to show that, for the minerals he considered, there was no need to invoke the presence of structurally altered zones at the surface of the solid phase to explain the observed pressure dependence. However, molal volume measurements over a wider temperature and concentration range will be needed in order to estimate the effects of pressure on reactions in the geothermal brines produced along saline spreading centres on the ocean floor. It is now becoming apparent that the reactions occurring at these sites may be responsible for fixing the concentration of many seawater components, and that the more spectacular deep-ocean vents may be surrounded by rich mineral deposits whose formation and stability require further investigation. □

Michael Whitfield is at the Marine Biological Association of the United Kingdom, Plymouth PL1 2PB.

1. UNESCO *Background Papers and Supporting Data on the International Equation of State for Sea Water 1980* (Paris, 1981).
2. Millero, F.J. *Geochim. cosmochim. Acta* **46**, 11 (1982).

0028-0836/82/240538-02$01.00

© 1982 Macmillan Journals Ltd

The impact, transmission and evolution of infectious diseases

from Robert M. May

OVER the past ten years and more, the World Health Organization (WHO) has increasingly placed its emphasis in the Third World on primary health care — paradigmatically exemplified by 'barefoot doctors' in villages — rather than on heroic technology and large hospitals[1]. I think this is very sensible, as many of the afflictions that combine to make life spans in the Third World markedly shorter than those in the developed countries can be countered by relatively inexpensive and simple measures; rehydration treatments for infants with diarrhoea or public health practices that help lower parasite burdens give better value for each dollar expended than do CAT scanners.

In designing programmes of primary health care, insights into the overall transmission and maintenance of infections within host populations are valuable. This fact has paradoxical aspects, for to some the 'flute music' (Simon Levin's phrase) of mathematical epidemiology must seem like the most baroque manifestation of Western cultural imperialism, even less relevant to village paramedics than is open heart surgery. Motivated partly by the idea that theoreticians and practitioners have interesting things to say to each other in this area, a recent Dahlem Conference* brought together ecologists, population biologists, epidemiologists, veterinarians and a variety of medical people to examine the overall population biology of infectious diseases. Taking several background papers as read, the conferees divided into four groups to focus on the impact of diseases on human and other animal populations; the factors influencing transmission and maintenance of infection; the possibilities for control or eradication of infections; and the broad co-evolution of hosts and pathogens.

Impact

As reviewed for the meeting by Holmes (University of Alberta), anecdotes abound about the devastation that diseases can inflict on natural populations of plants and animals. It is, however, much more difficult to assess the extent to which diseases are important in regulating such populations. For instance, in some species of wildfowl in North America some 80–90 per cent of the individuals not shot by hunters die of diseases each year, yet it is arguable that availability of breeding sites remains the primary factor regulating population density. The discussion made it plain that the various causes of deaths are hard to disentangle in most vertebrate populations, and left the impression that regulatory mechanisms may often depend on the interplay of several factors. Thus in those parts of Europe where rabies is now endemic in fox populations, the fox density may be significantly depressed below the disease-free level, which is set by territorial considerations; the present densities appear to depend on both the disease and territory size.

Holmes advanced the suggestion that invertebrate populations may more typically be regulated by infectious diseases and parasites than are vertebrate ones. Although some empirical evidence points in this direction, a systematic study is lacking. Theory suggests[2] that host populations are more easily regulated by infectious diseases if there is little or no acquired immunity (typically making the regulation of invertebrates, which lack such acquired immunity, easier than of vertebrates), and if the intrinsic rate of growth of the host population is not too large.

In any assessment of the impact of disease on natural populations, *Homo sapiens* must be considered as a special case. This is not out of respect for Bishop Wilberforce and his spiritual heirs, but because the patterns of explosive growth of human populations over the past century and more are without parallel elsewhere in the animal kingdom.

Changing life expectancies in human populations, particularly over the past two centuries in the developed world and since World War II in the developing world, are almost wholly due to reduced mortality from infectious diseases[3,4]. Although these diminished mortality rates and their demographic consequences are clear, the relative importance of the contributions made to them by better nutrition, improved hygiene, medical advances and other factors remains the subject of sharp and unresolved debate[3].

Transmission

Many biologists and medical people are fascinated by the intricate details that make each association between a host species and a virus, bacterium, protozoan or helminth unique. But underlying this richness and peculiarity of detail — which must be kept in sight as one moves towards control programmes — are some common themes.

A lot of discussion centred around the extent to which R_0, the 'intrinsic reproductive rate' of the parasite[5-8], can be used to characterize particular host–parasite associations. Calculation of R_0 itself involves classifying parasitic infections according to their overall dynamics rather than their conventional taxonomy. For microparasites, which broadly are those with high rates of direct reproduction in the definitive host (as typified by most viral and bacterial, and many protozoan, infections), R_0 is the average number of secondary infections produced when one infected host is introduced into an uninfected population. For macroparasites, which broadly are characterized by having no direct reproduction within the definitive host (as is typified by most helminth infections), R_0 is the average number of adult offspring established in 'second generation' hosts that one adult parasite produces in an essentially uninfected host population. In either case, R_0 depends both on the basic biology of the host–parasite interaction, and on the environmental and social factors that influence transmission rates.

An accurate estimate of R_0 for microparasites can be made if good serological studies are available. Failing this, rough estimates can be made from the approximate formula[6] $R_0 = 1 + (L/A)$ where L is the average life span of the human or other host, and A the average age at which infections are acquired when the disease is endemic. Estimation of R_0 for macroparasites is more difficult, and involves knowledge of the average parasite burdens, of the way in which parasites are distributed within the host population (almost invariably in a clumped or 'overdispersed' manner, rather than independently, randomly or uniformly), and of the way egg output and adult parasite survival depend on the number of parasites in a host.

In addition to attempting to estimate R_0 for a selection of microparasites and macroparasites, the conferees also catalogued some of the difficulties inherent in the concept. For one thing, not all infections lend themselves to ready classification in this binary scheme. For another, some microparasites have two or more qualitatively distinct categories of host. For example, while typhoid usually runs a relatively short course of infection from which the host either recovers or dies, a few hosts ('Typhoid Marys') will be asymptomatic yet will carry and can transmit the infection for a relatively long time, possibly even lifelong; it is not obvious which is more important in transmission, the much larger number of much shorter infections, or the small number of long and asymptomatic infections. Gonorrhoea and some other infections may be primarily maintained by a core of 'superspreaders'[9-12]. Such refinements can be embraced within the definition of R_0, but they complicate its estimation.

Control

Control is a term often used loosely to describe any reduction in the incidence of an infection. Clearly, eradication of the infection is a limiting case. Whatever the

*The Dahlem Conference on the 'Population Biology of Infectious Diseases' was held in Berlin on 15–19 March 1982. The proceedings will be published as a book by Springer Verlag, later this year.

Robert M. May is Class of 1877 Professor of Zoology at Princeton University, Princeton, New Jersey 08544.

0028-0836/82/240539-02$01.00

© 1982 Macmillan Journals Ltd

degree of control being attempted, the goal will tend to be relatively easier if R_0 is small. It was remarked by Molineaux (WHO, Geneva) that the WHO success in the global eradication of smallpox and the failure in efforts to eradicate malaria from large parts of Africa may be partly because for smallpox in developing countries R_0 is $\sim$ 3 or 4, while for *falciparum* malaria R_0 ranges from $\sim$ 50 to several hundreds.

The existence of reasonably safe and effective vaccines against many microparasitic infections provides the basis for attempts at control or even eradication. Again, knowledge of R_0 can give a rough idea of the prospects for such campaigns. Assuming the vaccine to be 100 per cent effective (which few are), eradication requires that the fraction, p, of the population protected by vaccination must exceed a critical value[6,7,13] $p>1-1/R_0$. For large R_0, this can require a forbiddingly high proportion of the population to be vaccinated. Thus, although the aetiologies of measles and smallpox are in many ways similar, with the infection always being apparent and running a relatively short course, the tentative estimate that R_0 is probably around 10–20 for measles in developing countries may make its global eradication much more difficult than was the case for smallpox with its R_0 of 3–4.

For macroparasites, there was discussion of the distinction between infection (having one or more parasite) and disease (having a parasite burden sufficiently high to produce illness). Especially if parasite burdens are highly overdispersed, as they usually are, relatively few hosts may carry burdens high enough to induce illness, even though essentially everyone is infected. There can thus, as emphasized by Warren[14] and others, be a real distinction between 'infection' and 'disease', and it may make sense to target control measures against the heavily infected hosts.

The concept of a 'breakpoint', first introduced by Macdonald[15] in his thinking about control of schistosomiasis, is an interesting one for many macroparasitic infections. If the parasite has a sexual stage in the definitive host, as many macroparasites do, then there can be two alternative stable states for the host–parasite association above the threshold at $R_0 = 1$: if the parasite population is present at densities above the 'breakpoint' level, the chances that female parasites in a given host will be mated is high, and the parasite population will be able to realise its intrinsic reproductive rate and establish itself; but if the parasite population is initially at a density below the breakpoint level, the chance for a given female to acquire a mate is low, and the infection will die out (even though R_0 for mated females exceeds unity). This notion of a breakpoint is obviously attractive. It offers the seductive hope that macroparasites could be eradicated by a one-time reduction below breakpoint densities, without the environmental or social changes that are needed if R_0 itself is to be reduced below unity. This contrasts with most other parasites, where, in the absence of a breakpoint phenomenon, unremitting vigilance is necessary for parasites with high R_0, even after the infection has been eradicated from a region; the basic biology is such that the infection will tend to re-establish itself if a few infected individuals are introduced. Reviewing recent work on this subject[16,17], the conferees sadly concluded that such breakpoint densities are probably too low to be of practical significance (mainly because the patterns of parasite clumping in hosts, mentioned above, tend to make mated pairs likely even at low density).

Throughout the week, it was often observed that many of the infections that excite attention in the West are of little significance in any global league table of diseases. For example, increasing incidence of various infections associated with homosexuals in large cities is attracting considerable attention, yet [as summarized for the conferees in Paul Mann's (Public Health Laboratory, Bath) clerihue] "Malaria/Is scarier/Than catching clap/From a travelling chap".

Evolution

The notion that 'successful' or 'well adapted' parasites are relatively harmless to their hosts is set forth as the received wisdom in most medical texts, and elsewhere. The idea is, at first sight, not unreasonable: all else being equal, it is to the advantage of both host and parasite for the parasite to inflict little damage. Moreover, a certain amount of anecdotal information supports this view. In his background paper, Allison (System Research, Palo Alto) observed that, for example, in regions of Africa where trypanosomiasis is endemic, indigenous ruminants suffer mild infections with insignificant morbidity, while domestic ruminants that have been bred for a long time in the region suffer more severely, with significant morbidity and mortality, and recently imported exotic ruminants suffer virulent infections which are usually fatal if untreated.

From a theoretical point of view, it would indeed appear that parasites evolve to be avirulent, provided that transmissibility and duration of infectiousness are entirely independent of virulence. This assumption, however, is not generally valid; the damage inflicted on their hosts by most viral, bacterial, protozoan and helminth parasites is often directly associated with the production of transmission stages. Once these complications are introduced into the theoretical models[18-20], it appears that many co-evolutionary paths are possible, depending on the details of the interplay between the virulence and the transmissibility of the parasite.

Considerable discussion centred around one unusually well documented example[21,22]. In the early 1950s, the myxoma virus was introduced into rabbit populations in Australia and England. At first the disease was highly virulent, but over the subsequent decade successively less virulent strains of the virus began to appear. Since the mid 1960s, the virus appears to have come to an equilibrium with its rabbit host (in both Australia and England), with the predominant strain of the virus being one of intermediate virulence. These data can be analysed to get an approximate estimate of the relationship between the virulence and the transmissibility of the various strains of the myxoma virus[20]; the nature of this relationship is such that strains with intermediate virulence have the largest R_0, and may thus be expected to predominate.

There was discussion of other cases where the evolutionary pressures on the parasite may make for its being highly virulent. The various baculoviruses, which kill their insect hosts and effectively turn them into masses of viral transmission stages, are one such example that may be amenable to an explicit analysis of the contending evolutionary forces.

The group concentrating on evolutionary aspects of host–parasite interactions thus concluded that, although parasite 'harmlessness' may characterize many old-established associations, neither *a priori* arguments nor empirical evidence points towards its being a general rule. Several interesting questions that seem ripe for mathematical or experimental investigation were identified.

In all, most of the participants found the conference helpful both in summarizing where things stand, and in identifying directions in which we might productively move. Bryan Clarke spoke for everyone when he said:

If you're teased by a sense of unease,
We hope that this volume will please,
That the work clears the murk
From the quirks that may lurk
In the laws of the cause of disease.

1. Mahler, H. *Scient. Am.* **243** (Sept.), 66–77 (1980).
2. Anderson, R.M. & May, R.M. *Nature* **280**, 361, 455 (1979).
3. McKeown, T. *The Role of Medice: Dream, Mirage or Nemesis* (Princeton University Press, 1979).
4. Wrigley, E.A. & Schofield, R.S. *The Population History of England, 1541–1871* (Harvard University Press, 1981).
5. Macdonald, G. *Trop. Dis. Bull.* **49**, 813 (1952).
6. Dietz, K. in *Epidemiology* (eds Ludwig, D. & Cooke, K.L.) 104 (SIAM, Philadelphia, 1975).
7. Anderson, R.M. & May, R.M. *Science* **215**, 1053 (1982).
8. Yorke, J.A., Nathanson, N., Pianigiani, G. & Martin, J. *Am. J. Epidemiol.* **109**, 103 (1979).
9. Yorke, J.A., Hethcote, H.W. & Nold, A. *J. Sex Trans. Dis.* **5**, 51 (1978).
10. Nold, A. *Math. Biosci.* **52**, 227 (1980).
11. Kemper, J.T. *Math. Biosci.* **52**, 227 (1980).
12. May, R.M. *Nature, News and Views* **291**, 376 (1981).
13. Smith, C.E.G. *Proc. R. Soc. Med.* **63**, 1181 (1970).
14. Warren, K.S. *A. Rev. pub. Hlth* **2**, 101 (1981).
15. Macdonald, G. *Trans. R. Soc. trop. Med. Hyg.* **59**, 489 (1965).
16. May, R.M. *Math. Biosci.* **35**, 301 (1977).
17. Bradley, D.J. & May, R.M. *Trans. R. Soc. trop. Med. Hyg.* **72**, 262 (1978).
18. Levin, S.A. & Pimentel, D. *Am. Nat.* **117**, 308 (1981).
19. Bremermann, H.J. *J. theor. Biol.* **87**, 671 (1980).
20. Anderson, R.M. & May, R.M. *Parasitology* (in the press).
21. Fenner, F. & Radcliffe, F.N. *Myxomatosis* (Cambridge University Press, 1966).
22. Ross, J. in *Animal Disease in Relation to Animal Conservation* (Symposium Zoological Society, London, in the press).

©1982Nature Publishing Group

REVIEW ARTICLE

The relationships of *Sivapithecus* and *Ramapithecus* and the evolution of the orang-utan

Peter Andrews* & J. E. Cronin†‡

* Department of Palaeontology, British Museum (Natural History), Cromwell Road, London SW7 5BD, UK
† Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138, USA

We review here the molecular data that bear on and provide a framework for interpreting hominoid relationships. Man is shown to be most closely related to chimpanzees and gorillas among extant hominoids, with the orang-utan more distantly related to them and the gibbons more distantly still. A fossil ape, Sivapithecus meteai, *shares several characters with the orang-utan and is thus probably related to it.* S. meteai *is part of the Middle Miocene* Sivapithecus–Ramapithecus *species complex, and if this group forms a valid clade then* Ramapithecus *must also be considered as being more closely related to the orang-utan than to man. The date of divergence of the orang-utan from the African apes and man is suggested by fossil and molecular evidence to be 10±3 Myr ago.*

THE Miocene epoch (25–5.5 Myr) was the period that heralded the appearance of mammalian faunas of modern appearance. It was also the period that saw the main radiations of the hominoid primates, culminating in forms that share many of the characteristics of the living apes and man[1,2]. The specific affinities within this complex of fossil and living hominoids are still, however, subjects of much controversy[1–3], and other than recent Pleistocene forms which are obviously related to the living orang-utan[4] and gibbons[5], and the Plio-Pleistocene forms of early hominids from East Africa[6,7], no fossil hominoid has been convincingly shown to be related uniquely to any of the extant hominoids.

Foremost in this controversy is the postulated hominid affinity of *Ramapithecus*, a form found now in middle to late Miocene deposits in East Africa[1,8], Turkey[9], China[10] and Indo-Pakistan[3,11,12]. This phyletic linkage has been much debated and has been both rejected and supported on morphological[11,13] and molecular grounds[14–16]. Central to understanding this problem is: (1) the relationship of the complex of species within the *Sivapithecus–Ramapithecus* network, and (2) the relationships among the living hominoids and in particular the status of the orang-utan. To a large extent the adaptation and relationship of the orang-utan have remained substantial enigmas[14,17,18], but they are clearly pivotal to our understanding of ape and hominoid evolution.

Recently, the relationships among the living species of hominoids have been greatly clarified by macromolecular comparisons[14,15,17,19]. Although not yet definitive, these comparisons have provided a new insight into the branching sequence or cladistic relationships of the living apes and the timing of the divergence between apes and man. This evidence will be reviewed here, with particular reference to the relationships of the orang-utan with the other hominoids. Changes in morphology, particularly of the teeth and face, will be correlated with the pattern of relationships thus established.

As regards the fossils, a new fossil hominoid specimen has been recently described which sheds a substantial amount of light on the *Sivapithecus–Ramapithecus* complex. *S. meteai*, from the late Miocene of Turkey[20], shows similarities in morphology and a possible phyletic relationship to the living orang-utan. *S. meteai* clearly shares broad similarities with other Miocene *Sivapithecus*[20] forms, which in turn are closely allied to *Ramapithecus*[11]. If this linkage is correct then one segment of the hominoid evolutionary tree can be rooted through a purely palaeontological approach.

Molecular affinities among the hominoids

Each organism carries within itself a history that is encoded in the genome. Given the appropriate techniques we can decipher this codex and elucidate phylogenetic history. Such macromolecular comparisons may yield insights into both the sequence and timing of ancient speciation events. The molecular approach has long been critical to the resolution of several controversies concerning phyletic links over a wide taxonomic scale, for example, in the case of taxa such as Aves, Amphibia, Reptilia and Mammalia, particularly Primates. Such comparisons are not limited to living forms, as tissue from a frozen mammoth has supported its close molecular relationship to the modern elephant[21], and skin from the recently extinct Tasmanian wolf[21] has confirmed its affinities with the dasyuroid group of marsupials.

With respect to hominoid evolution, substantial genetic data are now available which shed light on the evolution of humans and apes. Data from immunodiffusion, microcomplement fixation, radioimmunoassay, amino acid sequencing, electrophoresis, nucleic acid hybridization, nucleotide sequence and restriction endonuclease mapping, and cytogenetics, have shown a fairly consistent pattern of the branching sequence or cladistics of the hominoids, and working within this framework one can attempt to analyse the rate and nature of morphological evolution along the specific lineages.

The biomolecular evidence (Fig. 1) argues for three clear major cladogenic events in hominoid evolution: first, the separation of the gibbon lineage from that leading to the great apes and man: second, the subsequent divergence of the orang-utan lineage from that linking the African apes and man; and thirdly, evidence from molecular comparisons that *Homo*, *Pan* (chimpanzee) and *Gorilla* share a substantial lineage before their subsequent divergence. Evidence as to which of the two extant lineages shares a period of common ancestry to the exclusion of the third is at present equivocal; evidence exists which favours pairing of *Pan–Gorilla*, *Homo–Pan* and *Homo–Gorilla* (in order of greatest preference). The difficulty in determining the sequence of divergence is primarily a result of extremely short

‡ Present address: Department of Anthropology, University of California, Berkeley, California 94720, USA.

0028–0836/82/240541—06$01.00

© 1982 Macmillan Journals Ltd

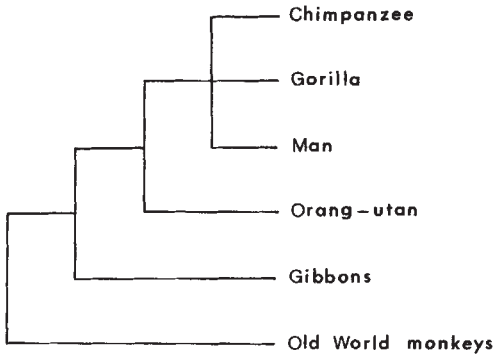


Fig. 1 A molecular cladogram of the Hominoidea incorporating data from immunological studies, electrophoretic analysis, protein sequencing, nucleic acid hybridization and restriction endonuclease analysis of mitochondrial DNA. The data are discussed in the text and in refs 22 and 23 in more detail. Molecular clock calibration of the various sequences of divergence (as calibrated in the text and in refs 15, 22 and 27) are (1) *Homo–Pan–Gorilla*, 5±15 Myr ago; (2) *Pongo–Homo, Pan, Gorilla*, 10±3 Myr ago; and (3) *Hylobates–Pongo, Homo, Pan, Gorilla*, 12±3 Myr ago.

lineages, in a temporal sense, and of the experimental error inherent in the techniques, and so far no data have conclusively demonstrated any unique shared derived molecular changes for any pair of the lineages. However, as interesting as this issue is, it is not central to the evidence for the common ancestry of the African apes and man and to the phyletic position of the orang-utan (Fig. 1).

Molecular data demonstrate convincingly the existence of the common lineage leading to the African apes and *Homo*. Along this common lineage (Fig. 1) numerous shared molecular changes have occurred. These involve derived amino acid sequence changes in the proteins albumin and transferrin, fibrinopeptides, carbonic anhydrase, haemoglobins, myoglobin, changes in mitochondrial DNA restriction sites, 1.5% base pair substitutions (involving $\sim 10^6$–10^7 base pair mutations) in nuclear DNA, and at least one locus with electrophoretic charge changes (refs 22, 23; Table 1). Specifically determined events and inferred molecular changes clearly document the existence of this lineage, and no molecular data reverse the position of *Pongo* and *Homo* shown in Fig. 1, although some data provide more reliable results than others. The conclusion, therefore, is that, in so far as the molecular evidence reveals phylogenetic relationships, the orang-utan is removed from unique association with the African apes and is the sister group to the clade comprising the African apes and man.

In addition to the branching pattern, the molecular data may also elucidate the timing of evolutionary events. It has been hypothesized that due to their apparent approximate regularity of change, molecules may serve as an evolutionary clock to date past divergence events[14,24], and this has been used to suggest that the human and African ape lineages separated as recently as 4–6 Myr ago. Calculations based on rates of nucleic acid and serum protein evolution[15] give dates ranging from ~8.5–14 Myr for the origin of the orang-utan lineage ($\bar{x} = 9.9 \pm 1.3$, $\gamma = 1.7$). An estimate of 10±3 Myr is thus quite in order, given the approximate nature of the molecular clock[25], and this agrees with the fossil evidence described here for the origin of the orang-utan lineage. Previously, the fossil evidence has appeared to indicate dates of ~15–20 Myr for this event, but the data given here suggest that the date should be close to or slightly earlier than 10–11 Myr. Precluding naive interpretations of the molecular clock[25], these results are quite compatible with each other[26,27].

Morphological affinities among the hominoids

The evidence from gross morphology generally agrees with the molecular evidence, except that it defines the relationships among the extant hominoids less precisely. The evidence must be viewed at various levels. First, there is the evidence linking the hominoids as a monophyletic group through such shared derived characters as the presence of a vermiform appendix and the loss of the tail. Second, there are the numerous characters of the postcranial and axial skeleton which link the great apes and man and exclude the gibbons[28], particularly, for example, the morphology of the wrist[29], the structure and development of the shoulder musculature[30] and the reduced lumbar region of the vertebral column[31]. Finally, there are the characters that provide evidence of pairing relationships within the great apes and man, and as these are the principal concern of this paper these characters will be examined in more detail.

Evidence exists which previously led to the grouping of the great apes as a single group containing the chimpanzee, gorilla and orang-utan. These all have a diploid number of chromosomes of 48 compared with 46 in man. They have many superficial similarities in skull morphology, such as the prognathism of the face or the enlarged nuchal area of the occiput[32], and in their postcranial morphology, particularly in their limb proportions and shoulder morphology which were thought to be functionally adapted to brachiation[32]. More recently, most of these characters have been shown to be either allometric consequences of increased body size[28,33,34] or primitive characters retained by the living great apes[35] and therefore having no relevance to the within-group relationships of the living great apes.

There is much stronger evidence linking the African apes, the chimpanzees and gorillas with each other, to the exclusion of the orang-utan, including the many anatomical specializations related to knuckle walking in chimpanzees and gorillas[36], and other characters such as the presence of the frontal sinus[37,38], a character which they also share with man. This evidence is in general agreement with the molecular data, and there is little morphological evidence that lends any support to alternative sets of relationships, for example, that the orang-utan is related more closely to gibbons[39], that it forms the sister group to the group comprising the gibbons, African apes and man[39], or that the orang-utan by contrast is more closely related to man than are the African apes[40].

The lower face of chimpanzees and gorillas has many points of similarity with the human face despite the gross differences

Table 1 Molecular comparisons among the great apes and *Homo*

Molecular comparisons	*Homo–Pongo*	*Pan/Gorilla–Pongo*	*Homo–Pan–Gorilla*
Fibrinopeptide sequence (no. of amino acid differences)[23]	2	2	0
DNA hybridization (% base pair mismatch)[40]	2.4	2.0	1.1
DNA hybridization (% base pair mismatch)[23]	—	4.5	2.4
Immunological distance (ID units)[14]	37	39	14
Immunodiffusion distance (antigenic distance)[16,23]	1.6	1.6	0.8
Electrophoresis: plasma proteins[14,27]	≫2	≫2	~1.6
cladistic analysis based on number of electrophoretic charge changes at 23 loci[22]	8	9	7
genetic distance value based on 23 loci[17]	0.349	0.300	0.367
Mitochondrial DNA comparisons using endonuclease restriction enzyme maps of site shared[19]	1*	2†	7‡

* *Homo* shares a specific site with *Pongo* but others do not; † *Pan* and *Gorilla* share a site in common with *Pongo* but *Homo* does not; ‡ *Homo* shares this number of sites with at least one of the species of *Pan* or *Gorilla* but not with *Pongo* (or *Hylobates*).

© 1982 Nature Publishing Group

Fig. 2 The maxilla of *S. meteai* showing the lower part of the face.

in facial height and the alveolar prognathism. Some of the characters shared between man, chimpanzees and gorillas are: the nasal aperture is broad; the subnasal plane is truncated and stepped down to the floor of the nasal cavity; the orbits are approximately square and often broader than high; the inter-orbital distance is broad; the infra-orbital foramina are usually three or less in number and are situated on or close to the zygomaticomaxillary suture; the zygomatic bone is usually curved and has a pronounced posterior slope; the zygomatic foramina (one or two) are small and are situated at or below the lower rim of the orbits; and the glabella is thickened (present in some fossil humans although not in modern man). The chimpanzee, gorilla and man also share the presence in the palate of small incisive foramina, large and oval-shaped greater palatine foramina, and large sphenopalatine fossae. Their teeth are basically similar in pattern except for the canines and premolars, and in particular the upper incisors lack a great size discrepancy between I^1 and I^2.

Many of the characters shared by man, chimpanzees and gorillas are present also in the gibbons. Where this is the case it is presumed that these characters are primitive for the Hominoidea, both because they are thus present in most extant hominoids and also because they are found in the most widely separated members of the group. (In other words, they are not present in the orang-utan, which separates the clades of gibbons and the African apes and man in Fig. 1.) In three of these characters, however, gibbons differ from the African apes and man, and these must be examined briefly to determine whether the primitive condition can be identified for them. Gibbons have a narrower oval-shaped nose, and from a consideration of other anthropoids, which generally have noses that are higher than broad, it seems likely that the primitive pattern for hominoids is a narrow oval-shaped nose as in gibbons. The infra-orbital foramina in gibbons are well removed from the zygomaticomaxillary suture, and again this is the usual condition in other anthropoids, thus it seems likely that this is the primitive condition in the hominoids. The condition of the glabella varies considerably in gibbons, from almost complete absence of thickening to fairly extensive thickening. This is true also of many other anthropoids, and it seems in fact to be size-related, therefore it is only possible to consider this character to be significant if glabellar thickening is either present on a small-bodied species or absent from a large one.

These characters have been summarized in a list (Table 2) which depicts the primitive condition in the Hominoidea. The two characters marked with an asterisk are present in gibbons, which thus retain the primitive condition, but not in the African apes and man, which thus share the derived conditions for these characters. All the other characters are shared between gibbons, African apes and man. One other character for which the polarity of evolutionary change is not yet understood concerns the presence of thick or thin enamel on the molars. The gibbons, chimpanzees and gorillas have thin enamel and man has thick enamel, and, on the face of it, thin enamel seems to be primitive for hominoids. Gibbons, chimpanzees and gorillas would thus have retained the primitive condition while man has developed the derived condition with thick enamel. Thick enamel, however, is a character that is shared with the orang-utan as well as with the fossil apes of the *Sivapithecus–Ramapithecus* group, and an understanding of the relationships between these two groups is essential to the understanding of human relationships.

Relationships of the orang-utan

The characters of the orang-utan differ in most respects from those listed in the previous section. Table 3 shows the characters of the orang-utan which it does and does not share with the fossil *S. meteai*[20] (Fig. 2). Other characters which it does not share with *S. meteai* could be added to the list, for example, the presence of wrinkling on the molar teeth in the orang-utan; but the absence of such clearly derived characters in the fossil hominoid need not be construed as evidence against an association between it and the orang-utan. Evidence for such an association comes from the list of shared characters, all of which (except where indicated) are interpreted as derived characters uniquely shared between *S. meteai* and the orang-utan. Where the fossil hominoid lacks the derived characters present in the orang-utan, this can be interpreted in terms of mosaic evolution, whereby some characters in an evolving lineage are developed before others. It would be more difficult to explain derived characters present in the fossil that are not present now in the orang-utan, but none have been observed.

No evidence exists as to whether the characters shared between *S. meteai* and the orang-utan are homologous or not. They may have arisen through convergent evolution rather than through common ancestry, but the latter is indicated by several sources. First, the characters shared between the orang-utan and *S. meteai* relate to separate functional complexes of the face and dentition. Some of the characters do inter-relate, such as the shape of the nose and the subnasal plane, or the nose and orbit shapes, or even possibly the shape of the nose and the narrow inter-orbital distance, but these characters relate directly neither to the characters of the palatine foramina, nor to two characters of the dentition, that is, the large difference in size between the central and lateral incisors and the thickness of enamel on the molar teeth. These combinations of characters can be seen as separate functional complexes that show similarities between the orang-utan and *S. meteai*, and if they are all the results of convergence it is difficult to envisage the circumstances that would lead to similarities in such disparate sets of characters occurring together in the fossil and extant forms.

Homology rather than convergence is also indicated by the metrical study of *S. meteai*[41], in which over a series of eight measurements of the palate and lower face, the extant hominoids showed a consistent and unified pattern differing

Table 2 Characters considered likely to be primitive for the Hominoidea

*Nasal aperture higher than broad, oval-shaped
Subnasal plane truncated
Subnasal plane stepped down to floor of nasal cavity
Orbits as broad as or broader than high
Inter-orbital distance broad.
Infra-orbital foramina few in number (≤3)
*Infra-orbital foramina well removed from the zygomaticomaxillary suture
Zygomatic bone curved and with strong posterior slope
Zygomatic foramina small
Zygomatic foramina 1–2 in number
Zygomatic foramina situated at or below the lower rim of the orbits
Glabella thickening which may occur on large individuals/species
Small incisive foramina
Large, oval-shaped greater palatine foramina
Upper incisors lacking large size discrepancy
Thin enamel on molars (see text)

This list is not exhaustive but concentrates on those characters that are known from the fossil record.

* These characters are present in gibbons but not in the African apes and man.

© 1982 Nature Publishing Group

markedly from that of other anthropoids. The pattern for *S. meteai* approximates fairly closely to the great ape pattern, demonstrating both that its affinities in most dimensions are with the great apes and that, seen as a pattern, each of these characters is consistent with the others and therefore more likely to be the result of shared ancestry rather than independently acquired convergences. As for the non-metrical data, the variation in some of the dimensions, for example, palate length and nasal width, are not significantly inter-correlated and by inference are not functionally related.

Even more tenuously, it can be observed that in certain other respects the face of *S. meteai* bears strong resemblances to that of the orang-utan. The size and conformation of the zygomatic region, for example, are similar in the two hominoids, although the degree of similarity is not sufficient to distinguish them absolutely from the other hominoids. The great bizygomatic breadth in *S. meteai* is most similar to the condition in the orang-utan, but it also falls within the range of variation for male gorillas[20]. The same is true of the relatively great depth of the zygomatic region and the pronounced premaxillary prognathism[20]. The deep zygomatic region is combined in *S. meteai* with an estimated face length (nasal height measured from the estimated position of nasion[20]) that is below the range for the gorillas but within the range for the orang-utan. Thus these characters combine to produce a face shape similar to that of the orang-utan and different from comparable-sized gorillas. However, as the fossil is incomplete, it is not yet possible to quantify these characters adequately and little reliance can be placed on them except as support for the evidence given earlier.

It can be concluded from the morphological evidence given here that the orang-utan possesses a number of derived characters not present in the other extant hominoids. This is consistent with the molecular evidence discussed in the previous section (Fig. 1). We have shown that the suite of characters that distinguishes the orang-utan from other hominoids is present also in a fossil ape, *S. meteai*, and as it is concluded that these similarities are probably homologous we have felt justified in adding the fossil hominoid to the original cladogram to produce Fig. 3.

Sivapithecus indicus from India and Pakistan[3,11] can also be linked with *S. meteai* and the orang-utan on the basis of a newly described skull belonging to this species[42]. This skull, which was described after this review was written, shares with the orang-utan all the characters listed in Table 3, and it also provides evidence on two additional characters of the eye region which are included in Table 3. This strongly reinforces the case for the relationship of *Sivapithecus* with the orang-utan[44].

Discussion

S. meteai has long been identified as part of the *Sivapithecus–Ramapithecus* complex[3,20]. Simons and Pilbeam[3] synonymized

Table 3 Characters shared between the orang-utan and *S. meteai*

*Nasal aperture higher than broad, oval-shaped
Subnasal plane smooth, not truncated
Subnasal plane continuous with floor of nasal cavity, not stepped
Orbits higher than broad—not known for *S. meteai* but known for *S. indicus*[42]
Inter-orbital distance very narrow
*Infra-orbital foramina few in number
*Infra-orbital foramina well removed from the zygomaticomaxillary suture
Zygomatic bone flattened, facing anteriorly
Zygomatic foramina relatively large
No glabellar thickening—not known for *S. meteai* but known for *S. indicus*[42]
Extremely small incisive foramina
Slit-like greater palatine foramina
Great size discrepancy between I^1 and I^2
Thick enamel on molars
†Zygomatic foramen multiple (single in *S. meteai*)
†Zygomatic foramen situated above the level of the lower rim of the orbit (below the orbit in *S. meteai*)

* Characters considered to be shared primitive characters; all the rest are considered to be shared derived characters.

†Characters that are present in the orang-utan but for which *S. meteai* retains the primitive condition.

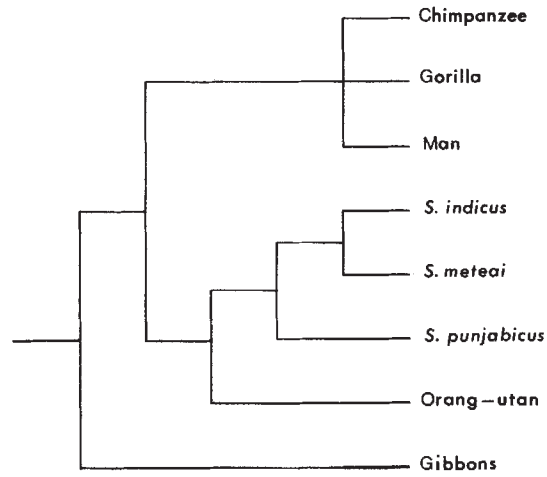


Fig. 3 Relationships of the extant and fossil Hominoidea. The relationships of the living species correspond to those in Fig. 1, and the fossils are shown as three species of *Sivapithecus* (including *Ramapithecus* as a junior synonym) which together constitute the sister group to the orang-utan.

it with *S. indicus*, but it was later reinstated as a separate species[20] on the basis of characters of the incisors, canines, premolars and posterior molars. In most of these characters it was presumed[20] that *S. meteai* possessed the derived condition with respect to the primitive hominoid morphotype and *S. indicus* the primitive condition. They both share, however, characters such as thickened enamel on the molar teeth, lower-crowned and more robust canines, and loss of molar cingula, which distinguish them from the primitive hominoid condition and indicate that the species presently assigned to *Sivapithecus* represent a clade, of which these are the distinguishing characters. As these characters are also shared by *Ramapithecus* it seems very probable that they are part of the same clade. This has been recognized implicitly by Pilbeam[11] in his combination of 'ramapithecines' and 'sivapithecines' in the family Ramapithecidae, and more explicitly by Greenfield[13] who dismissed even generic differences between them (although unfortunately he did so for the wrong reasons) and referred *Ramapithecus* to the genus *Sivapithecus*. This is certainly the simplest solution at present to a rather complicated situation, and although there is a risk that it is an over-simplification it does help to resolve many problems if the thick-enamelled Eurasian apes (excluding *Gigantopithecus*) are united in the genus *Sivapithecus*. This genus would then contain at least 3–4 species, which can be listed as follows with their geographical location and approximate ages:

S. indicus, India and Pakistan, middle to late Miocene, 11–8 Myr

S. punjabicus, India and Pakistan, middle to late Miocene, 11–8 Myr

Sivapithecus sp., China, late Miocene, 10–8 Myr

S. meteai, Turkey and Greece, late-middle Miocene, 11–10 Myr

Sivapithecus sivalensis[3] is of uncertain status, especially the type specimen of the species, and thus is omitted from the present discussion. The species of middle Miocene age assigned to *Sivapithecus* and *Ramapithecus*, *Sivapithecus darwini* from Turkey and Czechoslovakia[9] and *Sivapithecus africanus* and *Ramapithecus wickeri* from Kenya[8], are known only from extremely fragmentary dental remains and they lack the body parts that have been used in the present analysis to determine relationships. Therefore, these species also have been omitted from the present discussion.

Two consequences ensue from the recognition of the *Sivapithecus* group as a clade. First, if one member of the clade, that is, a species such as *S. meteai*, is accepted as being closely related to the orang-utan, then the other fossil species are also

© 1982 Nature Publishing Group

linked in the same association. This would apply, for example, to species of *Ramapithecus* if they are congeneric with *Sivapithecus*. Only if characters are found to be present in more complete specimens that exclude them from this association (and also of course from the genus *Sivapithecus*) can it be accepted that they do not form part of this clade. Following this line of argument, in conjunction with the molecular and morphological evidence for the relationship of the extant apes, it is evident that '*Ramapithecus*' may not be related exclusively to man but may be related to the orang-utan.

The case for the hominid affinities of ramapithecines has most recently been reviewed by Kay[43]. He also refers '*Ramapithecus*' to *Sivapithecus*, although not as a separate species but to the species *S. sivalensis*. He shows that the species of *Sivapithecus* share the following synapomorphies with man: (1) broad mandibular body; (2) low-crowned molars with thick enamel; (3) rather reduced canine size; (4) rather reduced canine sexual dimorphism; (5) buccolingually broad upper canines; and (6) tendency to enlarged P_3 metaconids. These are said to be characteristic of the whole *Sivapithecus* clade and that this is therefore the sister clade of man. We accept Kay's interpretation of the significance of characters 1 and 5, although mandibular robusticity varies considerably for the different species of *Sivapithecus*, but there are greater difficulties with the others. Two of the characters, the enlarged P_3 metaconids and the thick enamel of the molars, are present also in the orang-utan, and therefore provide equally good evidence for a *Sivapithecus–Pongo* relationship as for a *Sivapithecus–Homo* relationship. We do not accept Kay's conclusion that canine size is reduced in *Sivapithecus* for two reasons; first, his data show that the range of canine size relative to the first molar is equivalent to the middle of the range of extant great apes: in other words *Sivapithecus* canines are relatively bigger than female great ape canines as well as being relatively smaller than male canines. Secondly, it must be considered that *Sivapithecus* may have relatively large molars for its body size[11], and if this is the case the ratio of canine/molar size may be misleading. Finally, we also disagree with the significance attached to reduced sexual dimorphism, as low sexual dimorphism has not been shown convincingly to be a valid characteristic of *Sivapithecus*.

The characters linking *Sivapithecus* with man are thus reduced to two: robust mandibles and buccolingually broad upper canines, two characters that are almost certainly functionally related. These characters certainly seem to be synapomorphies for *Homo* and *Sivapithecus*, but the likelihood of their being the result of parallel evolution must be given credence in view of the more numerous and functionally independent derived characters of the nose, the orbits, the palatal foramina and the incisors that are shown here to be shared between *Sivapithecus* and *Pongo*. We therefore consider it more likely that *Sivapithecus* species, particularly *S. meteai* and *S. indicus* but by inference also *Sivapithecus* (formerly *Ramapithecus*) *punjabicus*, are more closely related to the orang-utan than to man.

The other consequence arising from the proposed relationship of *S. meteai* and the orang-utan is that it provides an approximate minimum divergence data between the orang-utan and its extant sister group, the African apes and man. *S. meteai* itself comes from deposits about 10–11 Myr old, but the time span of *Sivapithecus* as recognized here is of the order 8–11 Myr. This age corresponds to the date of 10 ± 3 Myr suggested by the molecular evidence and strongly indicates a minimum age of at least 10–11 Myr for the divergence between the orang-utan clade and the clade containing the African apes and man. The divergence between the members of the latter must be subsequent to this period, and the molecular evidence for a prolonged common ancestry between the African apes and man suggests that their eventual divergence was considerably later than the middle Miocene and perhaps as late as the end of the Miocene 5–6 Myr ago[15].

There is one other consequence of the relationships proposed here, which will be considered briefly. The classification of the cladistic groups proposed in Fig. 3 would differ greatly from many current schemes. The most recent one by Szalay and Delson[5], for example, combines the great apes in the tribe Pongini with parallel tribes for *Dryopithecus* (Dryopithecini) and *Sivapithecus* (Sugrivapithecini). These are combined in the subfamily Ponginae, and *Ramapithecus*, *Australopithecus* and man are classified in the subfamily Homininae. These subfamilies are then grouped in the family Hominidae. Classifications based on Fig. 3 would have to recognize the chimpanzee, gorilla and man as one clade with the orang-utan and *Sivapithecus* (and '*Ramapithecus*') as another; a possible solution would be to classify these clades at subfamily level, as do Szalay and Delson[5], but with Homininae for the African apes and man, and Ponginae for orang-utan and *Sivapithecus* (and '*Ramapithecus*'). The African apes and man would then be divided at the level of tribe. An alternative classification at the family level is as follows:

Hominoidea (Gray, 1825)
- Hylobatidae (Blyth, 1875)
  - *Hylobates* (Illiger, 1811)
- Pongidae (Elliot, 1913)
  - *Pongo* (Lacepede, 1799)
  - *Sivapithecus* (Pilgrim, 1910)
- Hominidae (Gray, 1825)
  - Gorillinae (Hurzeler, 1968)
    - *Pan* (Oken, 1816)
    - *Gorilla* (I. Geoffroy, 1852)
  - Homininae (Gray, 1825)
    - *Homo* (Linnaeus, 1758)
    - *Australopithecus* (Dart, 1925)

This change from traditional classifications has long been advocated by Goodman[23], and it is one that is necessary with the recognition that the African apes are more closely related to man than they are to the orang-utan.

We thank N. Boaz, L. Brunker, E. Delson, S. Gould, T. Harrison, J. Lowenstein, L. Martin, E. Meikle and C. Stringer for helpful comments. V. Sarich provided helpful discussion on the molecular section of this paper, and much of the morphological evidence is the product of an ongoing collaboration with C.P. Groves. The Wenner-Gren Foundation, the L.S.B. Leakey Foundation and the Milton Fund of Harvard University provided grants to J.E.C. in support of part of this work.

1. Pilbeam, D. R. *A. Rev. Anthrop.* **8,** 333–352 (1979).
2. Andrews, P. in *Aspects of Human Evolution*, 25–61 (Taylor & Francis, 1981).
3. Simons, E. L. & Pilbeam, D. R. *Folia Primatol.* **3,** 81–153 (1965).
4. Delson, E. in *Paleoanthropology in The Peoples Republic of China* 40–65 (Natn. Acad. Sci., Washington D.C., 1977).
5. Szalay, F. S. & Delson, E. *Evolutionary History of the Primates* (Academic, New York, 1979).
6. Howell, F. C. in *Evolution of African Mammals*, 154–248 (Harvard University Press, 1978).
7. Johanson, D. C. & White, T. D. *Science* **203,** 321–330 (1979).
8. Andrews, P. & Walker, A. *Human Origins*, 279–304 (Benjamin, Menlo Park, 1976).
9. Andrews, P. & Tobien, H. *Nature* **268,** 699–701 (1977).
10. Xu, Q. & Lu, Q. *Vertebr. palaeont.* **17,** 1–13 (1979).
11. Pilbeam, D. R. *et al. Nature* **270,** 689–695 (1977).
12. Simons, E. L. *Phil. Trans. R. Soc.* **292,** 21–41 (1981).
13. Greenfield, L. O. *Am. J. phys. Anthrop.* **50,** 527–546 (1979).
14. Sarich, V. M. & Cronin, J. E. in *Molecular Anthropology*, 141–170 (Plenum, New York, 1976).
15. Cronin, J. E. & Meikle, W. E. *Int. J. Primatol.* (in the press).
16. Goodman, M. in *Molecular Anthropology*, 321–353 (Plenum, New York, 1976).
17. Bruce, E. & Ayala, F. *Evolution* **33,** 1040–1056 (1979).
18. Smith, R. J. & Pilbeam, D. R. *Nature* **284,** 447–448 (1980).
19. Ferris, S. O., Wilson, A. C. & Brown, W. M. *Proc. natn. Acad. Sci. U.S.A.* **78,** 2432–2436 (1981).
20. Andrews, P. & Tekkaya, I. *Palaeontology* **23,** 85–95 (1980).
21. Lowenstein, J., Sarich, V. M. & Richardson, B. J. *Nature* **291,** 409–411 (1981).
22. Cronin, J. E. in *New Interpretations of Ape and Human Ancestry* (Plenum, New York, in the press).
23. Goodman, M. & Cronin, J. E. in *A History of Physical Anthropology 1930–1980* (Academic, New York, in the press).
24. Wilson, A. C., Carlson, S. S. & White, T. J. *A. Rev. Biochem.* **46,** 573–630 (1977).
25. Corruccini, R., Baba, M., Goodman, M., Ciochon, R. L. & Cronin, J. E. *Evolution* **34,** 1216–1219 (1980).
26. Zihlman, A. L., Cronin, J. E., Cramer, D. L. & Sarich, V. M. *Nature* **275,** 744–746 (1978).
27. Cronin, J. E. & Meikle, W. E. *Syst. Zool.* **28,** 259–269 (1979).
28. Andrews, P. & Groves, C. P. in *Gibbons and Siamang* Vol. 4, 167–218 (Karger, Basel, 1975).
29. Lewis, O. J. *Am. J. phys. Anthrop.* **30,** 251–268 (1969).

© 1982 Nature Publishing Group

30. Ashton, E. H. & Oxnard, C. E. *Trans. zool. Soc. Lond.* **29**, 553–650 (1963).
31. Schultz, A. H. *Primatologia* **4**, 5/1–5/66 (1961).
32. Le Gros Clark, W. E. *The Fossil Evidence for Human Evolution* (University of Chicago Press, 1955).
33. Aiello, L. in *Aspects of Human Evolution*, 63–97 (Taylor & Francis, 1981).
34. Biegert, J. & Maurer, R. *Folia Primatol.* **17**, 142–156 (1972).
35. Delson, E. & Andrews, P. in *Phylogeny of the Primates*, 405–446 (Plenum, New York, 1975).
36. Tuttle, R. H. *Am. J. phys. Anthrop.* **26**, 171–206 (1967).
37. Cave, A. J. E. & Haines, R. W. *J. Anat.* **74**, 493–523 (1940).
38. Cave, A. J. E. *Proc. zool. Soc. Lond.* **136**, 359–373 (1961).
39. Romero-Herrara, A. E., Lehmann, H., Castillo, O., Joysey, K. A. & Friday, A. E. *Nature* **261**, 162–164 (1976).
40. Benveniste, R. E. & Todaro, G. J. *Nature* **261**, 101–108 (1976).
41. McHenry, H. M., Andrews, P. & Corruccini, R. S. *Folia Primatol.* **33**, 241–252 (1980).
42. Pilbeam, D. R. *Nature* **295**, 232–234 (1982).
43. Kay, R. F. *Int. J. Primatol* (in the press).
44. Andrews, P. *Nature* **295**, 185–186 (1982).

ARTICLES

Ancient and modern slopes in the Tharsis region of Mars

P. J. Mouginis-Mark*, S. H. Zisk† & G. S. Downs‡

* Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA
† NEROC Haystack Observatory, Westford, Massachusetts 01886, USA
‡ Jet Propulsion Laboratory, Pasadena, California 91109, USA

The directions of lava flows in the Tharsis region of Mars are used to identify regional palaeo-slopes, vent areas and local topography. A comparison is made between these flow directions and the present day radar-measured topography; good agreement between these data sets indicates that little intra-regional tectonic deformation has occurred following the emplacement of the preserved lavas.

As the youngest centre of constructional volcanism on Mars[1,2], the Tharsis region has been the focus of considerable attention directed towards gaining an understanding of the geophysical[3,4], geochemical[5] and volcanological[6] evolution of the area. Debate has centred on the relative importance of regional tectonic uplift versus constructional volcanism as the primary mechanism for the production of the observed relief[7,8], together with the attendant implications that such hypotheses would hold for lithospheric structure and crustal–mantle petrology. Until recently, it was believed that the centre of the broad Tharsis dome was located close to the three Tharsis Ridge volcanoes, where the surface rose to nearly 10 km above the martian datum[9,10]. New Earth-based radar measurements have shown, however, that Tharsis is much lower on average, with the mean elevation only 2–4 km above the 6.1 mbar reference surface[11]. In addition, the centre of the dome is evidently located further to the south-east than previously believed and is now centred on Syria Planum[12].

Nevertheless, considerable local topography (including numerous volcanic constructs) exist within Tharsis by virtue of the many outpourings of lava which characterize the area. As a consequence of the changing lithospheric load associated with this volcanic activity, deformation of areas such as the base of Olympus Mons would be expected if the lithospheric thickness was $\lesssim$150 km (ref. 13). In an attempt to search for recent tectonic deformation of this kind within Tharsis, we compare here the ancient slope directions (deduced from the orientations of lava flows) with the present-day slopes (derived from Earth-based radar measurements). Due to their propensity to flow parallel to the direction of slope, the eruptions of lava provide information on the characteristics of the local topography at the time of flow emplacement. In addition, the travel directions and areal extent of the flows can be used to describe further the nature of the volcanic activity within Tharsis because they permit the different eruptive centres and their relative volumes of lava to be recognized.

Flow measurements

Individual lava flows within Tharsis were mapped either from high resolution (20–150 m per pixel) Viking Orbiter images or from the new medium resolution (200 m) orthophotomosaics prepared by the US Geological Survey. The positions of these lava flows are shown in Fig. 1, with the exception of those flows which are located on the high flanks of the shield volcanoes. A few flow units in addition to those mapped may also exist, particularly to the west of Ulysses Patera and north of Noctis Labyrinthus, where only poor photographic coverage exists due to atmospheric opacity; some lobate flow edges are visible in these regions, but the boundaries cannot be traced for sufficient distances to infer their direction of travel.

A total of 475 lava flows longer than 20 km have been mapped (Fig. 1). These flows have typical widths of 5–10 km on the higher slopes of the volcanoes and 15–35 km on the lower flanks. Additional lava flows that spread out over wide areas and have widths greatly in excess of 40 km were deliberately excluded from this analysis because they were deemed unsuitable for the determination of their flow directions. In certain instances, particularly to the west of Arsia Mons, it has also proved difficult to map the total extent of each flow, since many flow segments are either partially buried by superposed younger units or their boundaries have become indistinct as more subdued relief is encountered close to their vents. Thus, while the flows measured here rarely exceed 200 km in length, previous estimates of 300–400 km appear to be quite plausible[14,15], while flow length maxima of ~1,000 km may also be possible for fissure eruptions originating at low elevations. For the flows included within this analysis, independent estimates of the flow thicknesses made from shadow measurements[14] indicate values between 5–20 m for the steeper slopes and 20–65 m for the generally flatter surrounding terrain.

The map of Tharsis lava flows (Fig. 1) enables both the regional slopes and the individual eruptive centres at the time of flow emplacement to be recognized. In agreement with the general observations of Schaber *et al.*[14], our analysis shows that the major palaeo-slope within Tharsis was down towards the north-west, with a high point present between Arsia and Pavonis Montes and a low located near to Olympus Mons. From the distribution of the flows, the linear extent of this north-west slope was at least 2,500 km, and there were also downhill gradients which extended away from Pavonis and Ascraeus Montes towards the north-east for 2,000 km, south-east from Syria Planum for 1,200 km and westward from Arsia Mons for ~1,000 km. While most of the flows originated from the four major shield volcanoes[14–16], additional lava flows were

0028–0836/82/240546—05$01.00 © 1982 Macmillan Journals Ltd

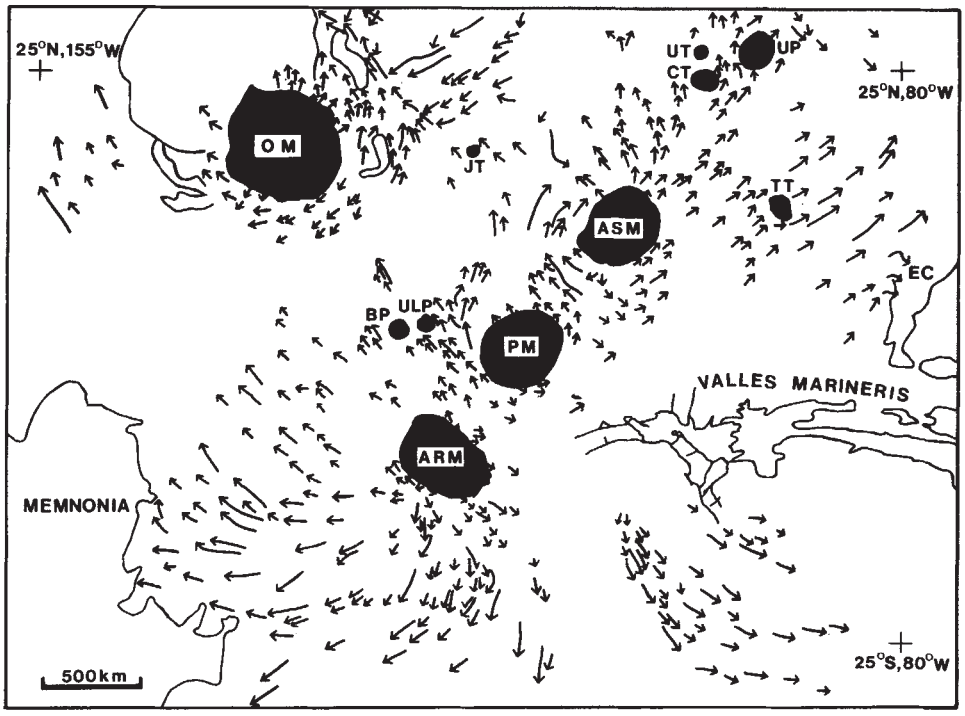


Fig. 1 Distribution of all lava flows included within this analysis. Length of arrow is equal to the length of the lava flow and shows the direction of travel. Volcanoes are shown in black where: ARM, Arsia Mons; ASM, Ascraeus Mons; BP, Biblis Patera; CT, Ceraunius Tholus; JT, Jovis Tholus; OM, Olympus Mons; TT, Tharsis Tholus; ULP, Ulysses Patera; UP, Uranius Patera; UT, Uranius Tholus. Also shown are the highland boundary in Memnonia (lower left), Valles Marineris, the Olympus Mons aureole (north-west of the volcano) and Echus Chasma (EC).

erupted in areas which lack obvious volcanic constructs (Fig. 1); Syria Planum and an area to the west of Ceraunius Fossae (25°N, 110°W) are two such examples. In these localities, several small cones of probable volcanic origin exist[17], but their significance as source regions for large volumes of magma had not previously been recorded. Lava flows >100 km in length originate from each of these two areas, indicating the existence of well developed fissure systems which were distinct from the conduits which supplied the magma to the major volcanoes.

Local deformations

In addition to illustrating the regional slopes within Tharsis, the orientations of the lava flows permit deformational features on a scale of a few hundred kilometres to be identified. Prominent amongst these is a circumferential depression, or peripheral trough, which surrounds the southern half of Olympus Mons[11]. Extending to distances of 300–350 km beyond the basal escarpment of the volcano, this depression has acted as a funnel to redirect lavas which consequently either flowed northward towards Arcadia Planitia or westward towards Amazonis Planitia. Segments of the aureole material to the east of Olympus Mons have been embayed and partially buried by lavas from both central Tharsis and Olympus Mons, while the much more extensive portions of the aureole to the north and west of the volcano, despite being down the regional slope[12], have remained unburied. From our investigation of the preserved record of volcanic activity, none of the lava flows erupted from Olympus Mons contributed to the lava pile which now constitutes central Tharsis (such as the area around Biblis and Ulysses Paterae); all the lava flows from the volcano travelled either to the north or west. This implies that the peripheral trough around Olympus Mons has been in existence for an extended period of time, rather than being a relatively recent response to crustal loading by the volcano after most of the lavas had been emplaced.

In contrast to Olympus Mons, the other three Tharsis shield volcanoes are noteworthy for the lack of preserved deformational features within the surrounding plains materials. A slight topographic rise ~330 km west of Ascraeus Mons, together with the tangential flow direction of lavas at the base of the volcano, suggest that a local depression may exist around this volcano, although subsequent infilling by lavas makes this identification uncertain. No similar evidence exists from the orientations of preserved lava flows for depressions around either Arsia or Pavonis Montes; flows from these two volcanoes travelled in almost radial directions from the summit calderas and adjacent fissure zones.

Only one example of surface deformation after lava flow emplacement can be found in the Tharsis region. To the north-east of the Olympus Mons peripheral trough, relatively old lava flows from the south-east have been partially buried and modified by more recent volcanism. In one instance, a sinuous lava channel was formed along a radial directed towards Olympus Mons and this lava channel cuts the older flows almost at right angles to the original flow direction (Fig. 2). Our interpretation of this dual flow direction is that after their emplacement, the older flows were tilted towards the south-west before the formation of the lava channel. Lithospheric loading of the area by the continuing construction of Olympus Mons could be the cause of this tectonic movement.

Present-day topography

Earth-based radar measurements of elevations on Mars have recently been referenced to the 6.1-mbar pressure surface[11].

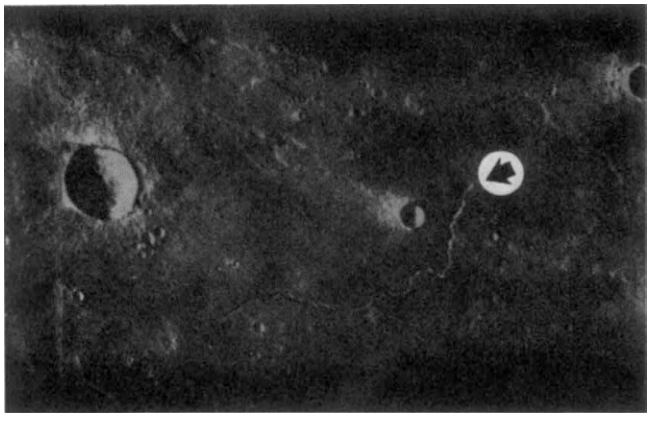


Fig. 2 North-east of Olympus Mons, a sinuous lava channel (source arrowed) cuts older lava flows which came from the south-east. The change in flow direction between these two eruptions is attributed to lithospheric loading associated with the construction of Olympus Mons. North is towards the top of image, which is centred at 29.8° N, 130.0° W. Image width is 60 km; Viking Orbiter frame 623A46.

©1982Nature Publishing Group

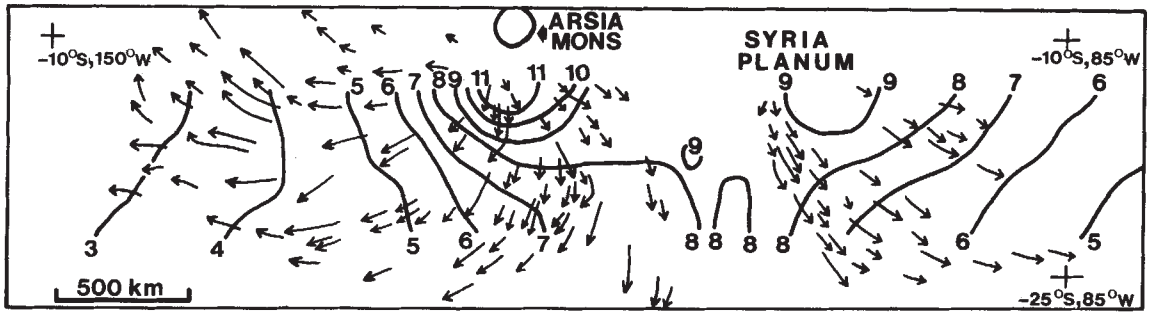


Fig. 3 Comparison of radar-derived topography and lava flow directions for the southern portion of the study area. Radar data give contours in 1-km intervals and are referenced to the 6.1-mbar pressure surface of Mars. Contours were hand drawn (by P.M.M.) from elevation measurements made in 1971 and 1973 between 14° and 21° S.

These data provide height information for the southern flanks of Arsia Mons and Syria Planum for latitudes 14–21°S. Data resolution for this area is ~10 km in longitude, 80 km in latitude and 100 m in altitude[18]. Using this radar-derived topography, it is therefore possible to compare the present day regional gradients with the observed flow directions (palaeo-slopes) to search for regional deformations which may have occurred after the cessation of the volcanic activity. Such tectonic deformation would be expected to be present if the lithosphere beneath Tharsis was thinner than ~150–180 km (refs 13, 19) due to the changing load associated with the protracted eruption of large volumes of lava[20,21].

The following comparison of these present and past slopes within southern Tharsis (Fig. 3) points to a strong agreement between the two data sets. The radar elevations corroborate that a regional slope extends downward from Arsia Mons towards Memnonia; the total change in elevation amounts today to >8 km. A slight (500–1,000 m) high is associated with the Claritas Fossae fractures, while Syria Planum represents the centre of a regional dome located at the western end of Valles Marineris[12]. Almost all the mapped lava flows cross the present day contour lines within about 10° of the local gradient (Fig. 3). Considering that the aforementioned resolution of the radar data provides only the general characteristics of the local topography, there is also good agreement between the regional slopes and the paths of the flows (such as those associated with the 'bending' of the flows south of Arsia Mons between latitudes 17°–22°, longitudes 110°–122°W). Essentially all of the observed lava flows are seen to travel generally down the present topographic gradient for their entire length. Close to the summit of Arsia Mons, however, where a parasitic shield is thought to be located[22], some flows do not follow the maximum slope. This may, on the other hand, be a consequence of the shallowness of the slopes (see below) and the relative thickness of the flows (~25 m (ref. 14)), which could have funnelled the younger flows around pre-existing units in a manner comparable to the diversion of certain terrestrial eruptions (ref. 23, p. 421).

Although the radar data set lacks sufficient areal coverage to compile a complete topographic map, scattered elevation values for parts of northern Tharsis are available[11] and permit the determination of the present-day slopes north of the equator. We have measured 14 slopes within Tharsis (all of which parallel the lava flow directions for distances of 150–1,250 km) using the radar data (Fig. 4). In each case, it is evident from the slope estimates (Table 1) that the lava flows travelled along paths that are consistent with the modern down-slope directions. All the radar-measured slopes seem to be very shallow, ranging from 0.75° for the southern flanks of Arsia Mons (profile L–L') to 0.04° for a section of the Olympus Mons peripheral trough (profile C–C'). These slopes are, however, only best estimates from the radar data because each elevation estimate is an average of several measurements made within a 2° longitude sample bin[11]. Typically, 7–13 measurements are included in each sample, and the resulting standard errors for the averages range from 70 to 770 m (Table 1). Over the measured horizontal distances considered here, the maximum error in the slope determination would be about ±0.17°, while typical errors would be <±0.1°. Under these circumstances, only in the case of profile C–C' could the down-slope direction of the profile change; although the absolute slopes would be different, all the remaining modern gradients are consistent with a down-slope movement of the lava flows even allowing for the resolution of the radar system.

Such a close match between the ancient and modern regional slopes indicates appreciable lithospheric stability within Tharsis. Although the absolute ages of the lava flows are not known, at least 14 distinct eruptive sequences of flows have been identified in this area[14] and these surface units have large differences in their superposed crater densities[21]. Crater counts give number densities ranging from 1.22×10^{-3} to 9.08×10^{-5} craters >1 km diameter per km^2 respectively for the Tharsis Montes unit 3 (tm_3) and Olympus Plains (op) units of Scott and Tanaka[21], suggesting a wide age range. It appears, therefore, that there was relatively little tilting of geological units during (or after) the extended period of Tharsis volcanism represented by the currently exposed lava flows. This in turn implies that the lithospheric structure beneath Tharsis was rigid enough to support the changing load associated with volcano construction, suggesting that the lithosphere may indeed have been thicker than ~150–180 km (refs 13, 19, 20).

Volcanology

In addition to the slope information, subtle differences in the style of Tharsis volcanism are also indicated by the distribution of the observed lava flows. From the large number of flows located between the three Tharsis Ridge volcanoes (Fig. 1), it

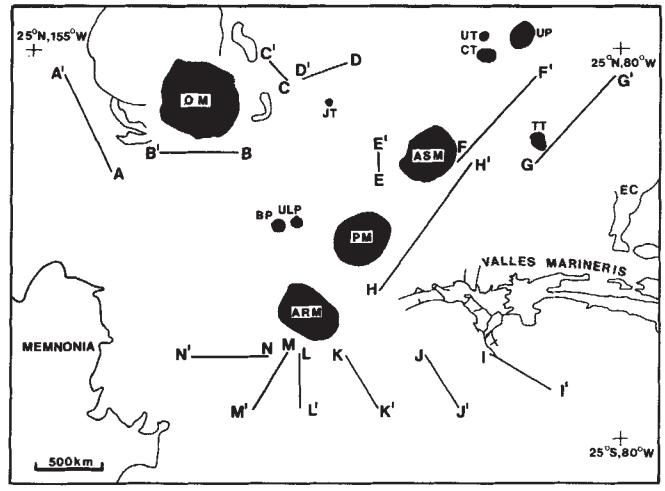


Fig. 4 Location of the surface slopes measured from Earth-based radar data (see Table 1). Each slope is coincident with the lava flow directions shown in Fig. 1, with the downslope measurement denoted by the superscript bar. Base map abbreviations are the Hsame as those in Fig. 1.

© 1982 Nature Publishing Group

Table 1 Radar-derived measurements of modern slopes in Tharsis region

	Up-slope (unprimed)					Down-slope (primed)							
	Lat. (deg)	Long. (deg)	Altitude* (km)	S	N	Lat. (deg)	Long. (deg)	Altitude* (km)	S	N	Distance† (km)	Height‡ (km)	Slope§ (deg)
A–A′	9.41	144–146	−0.49	0.15	9	22.00	150–152	−2.02	0.09	12	837	1.53	0.10±0.02
B–B′	12.11	128–130	2.22	0.10	11	12.11	138–140	0.07	0.18	9	600	2.15	0.20±0.03
C–C′	21.58	122–124	2.35	0.07	12	23.15	124–126	2.25	0.09	12	152	0.10	0.04±0.07
D–D′	23.15	114–116	4.45	0.14	12	21.58	120–122	2.65	0.13	13	372	1.80	0.28±0.04
E–E′	9.61	110–112	4.43	0.23	11	12.11	110–112	3.65	0.22	7	150	0.78	0.30±0.17
F–F′	10.85	100–102	4.90	0.25	3	21.78	90–92	3.47	0.08	9	889	1.43	0.09±0.02
G–G′	10.98	90–92	3.55	0.15	8	21.78	80–82	1.76	0.10	13	883	1.79	0.12±0.02
H–H′	−6.10	110–112	8.62	0.09	2	10.98	98–100	4.64	0.13	5	1252	3.98	0.18±0.01
I–I′	−14.47	96–98	8.74	0.10	13	−18.99	88–90	6.04	0.10	12	551	2.70	0.28±0.02
J–J′	−14.66	104–106	8.29	0.30	13	−21.11	100–102	7.63	0.10	12	455	0.66	0.09±0.05
K–K′	−14.47	114–116	8.74	0.05	12	−21.11	110–112	7.44	0.08	13	455	1.30	0.16±0.02
L–L′	−14.47	120–122	11.75	0.08	13	−21.11	120–122	6.69	0.12	12	386	5.06	0.75±0.03
M–M′	−14.47	122–124	10.98	0.57	12	−21.23	126–128	5.07	0.12	13	471	5.91	0.72±0.08
N–N′	−14.47	124–126	8.29	0.77	13	−14.46	134–136	4.32	0.13	12	600	3.97	0.38±0.09

* Altitudes are expressed in kilometres above the 6.1-mbar pressure surface. 1 σ (S) values from the mean elevation and the number of data points (N) in the two-degree sample bin are also given.

† Straight-line distances are given from the middle of the up- and down-slope points.

‡ Height difference between the mean up- and down-slope altitudes.

§ Surface slopes, estimated errors for these slope values are derived from the 1σ values in the altitude measurements.

is evident that large volumes of magma were also erupted at these localities from saddle fissures (in addition to the magma erupted from the summit calderas). Furthermore, a change in eruptive style from true shield volcanism (most of the magma erupted from near-summit vents) to flank activity (most of the magma finding egress at the saddle crests) evidently took place on both Ascraeus and Pavonis Montes late in their evolution, because no flows from their summit calderas extend across the saddle flows. Eruptions from Arsia Mons also display a similar change in vent location, with large eruptions of magma from the parasitic cone[22] postdating the near-summit activity. This sequence of vent relocations for these volcanoes is at least qualitatively consistent with the inferred history of the calderas of the Tharsis volcanoes[24].

All the small volcanic constructs of the Tharsis region (Biblis, Ulysses and Uranius Paterae and Tharsis, Uranius, Jovis and Ceraunius Tholi) show evidence of partial burial by lava flows which conform to the regional slopes identified in this analysis[25]. The absence of minor deviations from a straightline path, which most of the flows follow, indicates that there are no additional topographic rises within central Tharsis which could be other low volcanoes. Thus, it seems that the entire sequence of lava plains in this region[21] was constructed by a few major volcanic centres (not all of which were associated with the large shields); eruptions of magma were evidently confined to large, frequently reactivated fissure systems, rather than many small widely dispersed vents, during the period of activity presently exposed at the surface.

Comparing the size and number of the flows originating from Arsia Mons (42 > 100 km long) and Olympus Mons (3 > 100 km long) indicates that it was likely that these two volcanoes had appreciably different magma supply rates from their respective source regions. If it is assumed that these martian shields had eruption characteristics comparable to terrestrial volcanoes (that is, lava flow length was related to magma effusion rate[26]), the longer and more numerous flows from Arsia Mons would imply higher effusion rates than for Olympus Mons. Variations in effusion rate are also likely for different parts of Olympus Mons, because relatively long flows are only common on the eastern and southern flanks of the volcano while the aureole material to the north-west is almost devoid of embaying lavas. Thus, the traditional view that Olympus Mons is the largest volcano in the Solar System[1] may not be correct, because Arsia Mons not only has a larger caldera, but from this analysis also appears to have higher magma effusion rates and represents a topographic high that dominates the surrounding region for radial distances >1,000 km in any direction from its caldera. In comparison to both Arsia and Olympus Montes, the mass eruption rates for individual flows from Pavonis (seven flows longer than 100 km) and Ascraeus (nine flows longer than 100 km) were probably intermediate between these two extremes.

Conclusions

The use of lava flow directions as palaeo-slope indicators and Earth-based radar measurements of present topography demonstrate that little deformation of the Tharsis region has occurred over an extended period of time. Very few deformational features attributable to subsidence around the volcanoes can be recognized from the flow directions; a peripheral trough surrounds Olympus Mons and there is some evidence for a similar feature to the west of Ascraeus Mons. Only one example of surface tilting after lava emplacement has been found (to the north-east of Olympus Mons), suggesting that the increasing lithospheric load due to progressive buildup of the lava pile had little influence on local topography. It is therefore postulated that early in the evolution of Olympus Mons, the lithosphere was <150 km thick[13], thereby permitting the formation of the peripheral trough. For much of the period in which the preserved volcanic activity took place, however, lithospheric thickness has evidently been >150–180 km (refs 13, 19), as evidenced by the lack of local or regional deformation between the time of these eruptions and the present day.

Although regional slopes are very shallow today, if they truly reflect the gradients present at the time of volcanic activity, they nevertheless controlled lava flow directions over distances in excess of 1,000 km. In this respect, the rise associated with Syria Planum[12] has controlled the flow directions of many of the Tharsis lavas, while Claritas Fossae was also a local high at the time that Arsia Mons was active. It is evident that the style of volcanic activity was quite diverse within the region, with fissure eruptions from the saddles between the Tharsis Ridge volcanoes contributing large volumes of lava to the surrounding plains. In this connection, it is proposed that Arsia Mons, with its extensive lava flows and great relief, should supersede Olympus Mons for the title of the largest known volcano in the Solar System.

We thank Jim Head for his comments. This research was supported by NASA grant NER 40-002-088 of the Planetary Geology Program.

Received 18 February; accepted 1 April 1982.

1. Mutch, T. A., Arvidson, R. E., Head, J. W., Jones, K. L. & Saunders, R. S. *The Geology of Mars* (Princeton University Press, 1976).
2. Scott, D. H. & Carr, M. H. *Geologic Map of Mars* (U.S. Geological Survey Miscellaneous Map I-1083, 1978).
3. Sleep, N. H. & Phillips, R. J. *Geophys. Res. Lett.* **6**, 803–806 (1979).
4. Phillips, R. J. & Lambeck, K. *Rev. Geophys. Space Phys.* **18**, 27–76 (1980).
5. Finnerty, A. A. & Phillips, R. J. *Abstr. 3rd int. Colloq. Mars* 77–79 (Contr. 441, Lunar and Planetary Institute, Houston 1981).
6. Plescia, J. B. & Saunders, R. S. *Proc. 10th Lunar planet. Sci. Conf.* 2841–2859 (1979).

© 1982 Nature Publishing Group

7. Plescia, J. B. & Saunders, R. S. *Proc. 11th Lunar planet. Sci. Conf.* 2423–2436 (1980).
8. Solomon, S. C. & Head, J. W. *Lunar planet. Sci.* **11**, 1063–1065 (1980).
9. Christensen, E. J. *J. geophys. Res.* **80**, 2909–2913 (1975).
10. *Topographic Map of Mars* (U.S. Geological Survey Miscellaneous Map I-961, 1976).
11. Downs, G. S., Mouginis-Mark, P. J., Zisk, S. H. & Thompson, T. W. *J. geophys. Res.* (in the press).
12. Downs, G. S., Mouginis-Mark, P. J., Zisk, S. H. & Thompson, T. W. *Lunar planet. Sci.* **13**, 182–183 (1982).
13. Thurber, C. H. & Toksöz, M. N. *Geophys. Res. Lett.* **5**, 977–980 (1978).
14. Schaber, G. G., Horstman, K. C. & Dial, A. L. *Proc. 9th Lunar planet. Sci. Conf.* 3433–3458 (1978).
15. Carr, M. H., Greeley, R., Blasius, K. R., Guest, J. E. & Murray, J. B. *J. geophys. Res.* **82**, 3985–4015 (1977).
16. Crumpler, L. S. & Aubele, J. C. *Icarus* **34**, 496–511 (1978).
17. Moore, H. J. & Hodges, C. A. *NASA-TM 82385* 266–268 (1980).
18. Downs, G. S., Reichley, P. E. & Green, R. R. *Icarus* **26**, 276–312 (1975).
19. Turcotte, D. L., Willemann, R. J., Haxby, W. F. & Norberry, J. *J. geophys. Res.* **86**, 3951–3959 (1981).
20. Comer, R. P., Solomon, S. C. & Head, J. W. *Lunar planet. Sci.* **11**, 171–173 (1980).
21. Scott, D. H. & Tanaka, K. L. *Icarus* **45**, 304–319 (1981).
22. Roth, L. E., Downs, G. S., Saunders, R. S. & Schubert, G. *Icarus* **42**, 287–316 (1980).
23. Macdonald, G. A. *Volcanoes* (Prentice-Hall, New Jersey, 1972).
24. Mouginis-Mark, P. J. *Proc. 12th Lunar planet. Sci. Conf.* 1431–1447 (1981).
25. Pike, R. J. & Clow, G. D. *Abstr. 3rd int. Colloq. Mars* 199–201 (Contr. 441, Lunar and Planetary Institute, Houston, 1981).
26. Walker, G. P. L. *Phil. Trans. R. Soc.* **A274**, 107–118 (1973).

A model of the Antarctic Ice Sheet

J. Oerlemans

Institute for Meteorology and Oceanography, University of Utrecht, Princetonplein 5, Utrecht 3508 TA, The Netherlands

Numerical modelling of ice sheets and glaciers has become a useful tool in glaciological research. A model described here deals with the vertical mean ice velocity, is time dependent, computes bedrock adjustment and uses an empirical diagnostic relationship to derive the distribution of ice thickness in ice shelves. The rate of snowfall and ice/snow melt depends on the (prescribed) sea-level temperature, surface slope, elevation and distance to open water. The model is able to reproduce the major features of the Antarctic Ice Sheet. When it is run to a steady state for present climatic conditions, the main difference with the present ice sheet is that the shallow parts of the Weddell Sea become covered by grounded ice.

TO understand ice-sheet histories from proxy data, and to see whether they are physically plausible, numerical ice-sheet models can have an important role. Although the theory of ice flow is formulated reasonably well (this does not apply to what happens at the ice–bedrock interface), modelling of large ice sheets requires a pragmatic approach. It is not possible to solve the full stress–strain rate relationships, together with the thermodynamic equation, for the whole Antarctic Ice Sheet, for example. This certainly applies if one wants to carry out integrations over 100,000 yr or so.

In recent years, ice-sheet models based on a flow law for the vertical mean ice velocity have become popular[1–4]; they have been used mainly for palaeoclimatic studies, and in complexity come between perfectly plastic models[5] and three-dimensional models[6]. In these models temperature is not calculated, so constant bulk flow parameters have to be used. In spite of the simplifications involved, the vertically integrated models capture most of the characteristics of large ice sheets.

Based on these observations, I have developed a vertically integrated model of the Antarctic Ice Sheet. In fact the Antarctic Ice Sheet is an excellent test case for the 'vertical mean approach', because it has a very irregular bedrock topography and subtle grounding line dynamics at some places, and its present-day physical characteristics are well documented[7].

The first version of the present model was used to investigate how sensitive the Antarctic Ice Sheet is to changes in the ice accumulation rate[2]. However, this version did not include bedrock adjustment and ice shelves, and could only be used for short integrations.

The second version, results of which are presented here, has much more internal freedom. The response of the bedrock topography to a varying ice load is computed, and a diagnostic relation for ice shelves (though simple) is included. A varying bedrock topography introduces the problem of initial conditions. It is not very well known how close the present bedrock topography is to isostatic equilibrium. A 100-m error in the equilibrium bedrock elevation will be insignificant in the central parts of the East Antarctic Ice Sheet, but is very important in such marginal regions as the Ross Sea.

Model description

The calculation of ice flow is based on a flow law relating vertical-mean horizontal ice velocity u to basal shear stress τ_b

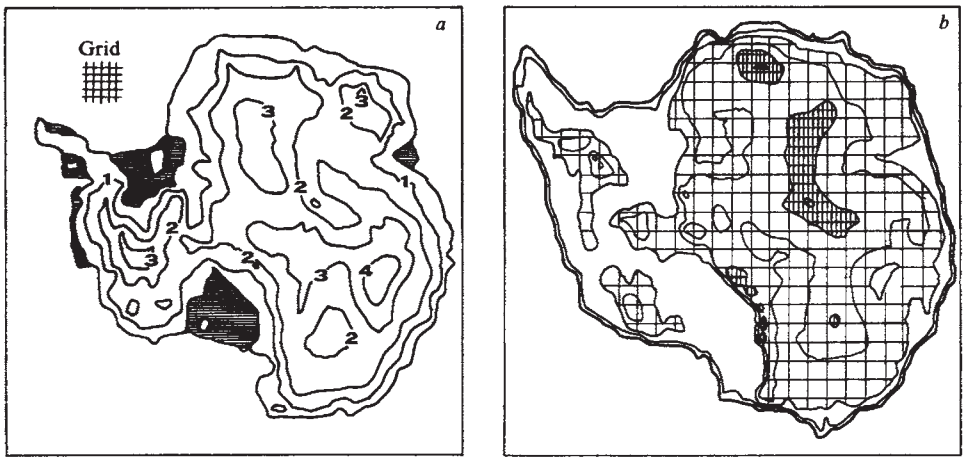


Fig. 1 *a*, Observed ice thickness as resolved on the model grid (part of which is shown in the upper left corner). Contour interval is 1 km. The major ice shelves are indicated by shading. *b*, Equilibrium bedrock topography in the case of no ice load. Contour interval is 1 km. The lowest contour level is 2 km below present sea level. Light shading indicates bedrock above sea level, heavy shading bedrock above the 2 km level. All input data are from ref. 10.

0028–0836/82/240550—04$01.00

© 1982 Macmillan Journals Ltd

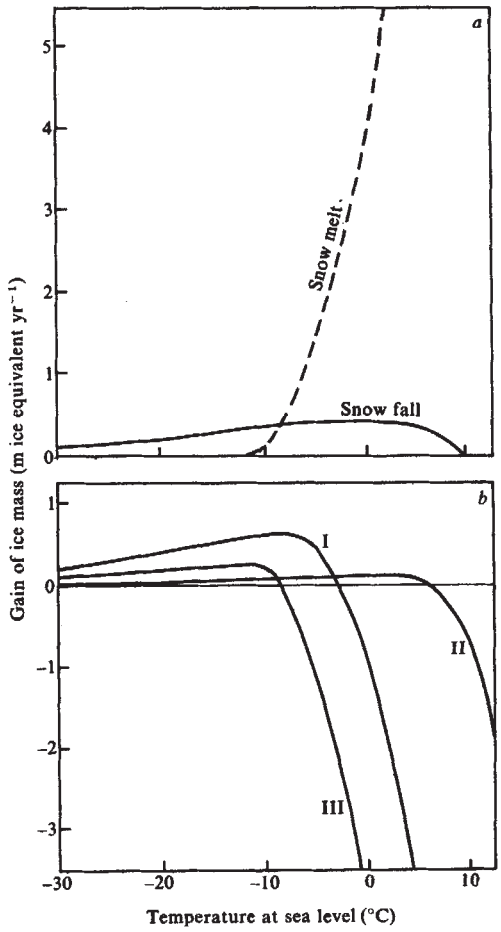


Fig. 2 Formulation of the mass balance, or ice accumulation rate, for the ice-sheet model. *a*, Dependence of snow/ice melt and snowfall on sea-level temperature for zero surface elevation and zero surface slope. *b*, The mass balance for different local conditions: I, a surface slope of 0.02 and elevation of 500 m, at the edge of the continent; II and III, the interior of the continent for elevations of 2,000 m and 0 m respectively.

(ref. 8). This law is of the Glen-type: $u = k\tau_b^m$, where k and m are flow parameters. The expression for u can be generalized to the two-dimensional case (involving assumptions that will not be discussed here), yielding the vertically integrated mass flow vector **M**:

$$\mathbf{M} = KH^{m+1}[\nabla H' \cdot \nabla H']^{(m-1)/2}\nabla H' \qquad (1)$$

where H is ice thickness, ∇ is the two-dimensional (horizontal) gradient operator and $H' = H + h$ is the surface elevation (h is the bedrock elevation with respect to sea level). So ice flows down the surface slope $\nabla H'$, and the rate at which this happens increases strongly with both ice thickness and surface slope. The evolution of the ice sheet is determined by the conservation of ice mass, that is

$$\partial H/\partial t = -\nabla \cdot \mathbf{M} + G \qquad (2)$$

G is the ice accumulation rate at the surface, expressed in units of ice depth per unit of time.

The response of the bedrock to the ice load is computed from:

$$\partial h/\partial t = -(h - h_0 + H/3.2)/T \qquad (3)$$

where h_0 is the equilibrium bedrock height without ice load. Equation (3) formulates a damped return, with time scale T, to local isostatic equilibrium. For loads with a horizontal length scale of more than a few hundred kilometres, equation (3) forms a good approximation[9]. For smaller loads, the flexural rigidity of the lithosphere becomes important. So, in general, equation (3) should work reasonably well, but for very small ice sheets or near the edge of a large ice sheet errors may occur (order of magnitude: 50 m).

Given the present bedrock topography and the distribution of ice thickness (Fig. 1*a*), and assuming that the present state of Antarctica is close to equilibrium, h_0 can be computed. Using data from the Russian Atlas of Antarctica[10], resolved on a grid with a spacing of 100 km, yields a no-ice bedrock topography h_0 as shown in Fig. 1b. This clearly shows how little of the bedrock beneath the present West Antarctic Ice Sheet is actually above sea level, even if the ice load is removed. Because we cannot be sure that the Antarctic Ice Sheet is close to a steady state, h_0 may be subject to errors. This applies in particular to the shallow Ross and Weddell seas, which were probably captured by grounded ice during the last ice age (20,000 yr ago). So it is possible that in these areas uplifting is still in progress[11]. Nevertheless, the bedrock topography displayed in Fig. 1*b* will serve as the initial condition for h, simply because a sound alternative is not available.

Ice shelves have to be taken into account to let the position of the grounding line be determined by the model. A prognostic equation for ice shelf growth/decay would not be feasible, because it would require a very small time step in the computations. Instead a diagnostic procedure is included. The thickness of floating ice at some point is calculated from the degree of enclosure (which is between 0° and 360°) by grounded ice, weighted with the distance to the grounding line and the ice thickness at the grounding line. In the present model only four directions (0°, 90°, 180° and 270°) are checked for each grid point to estimate the enclosure. An additional condition is that the ice shelf cannot be maintained if its thickness drops below 250 m. For an optimum choice of the proportionality constants appearing in this formulation, the calculated ice shelves for the

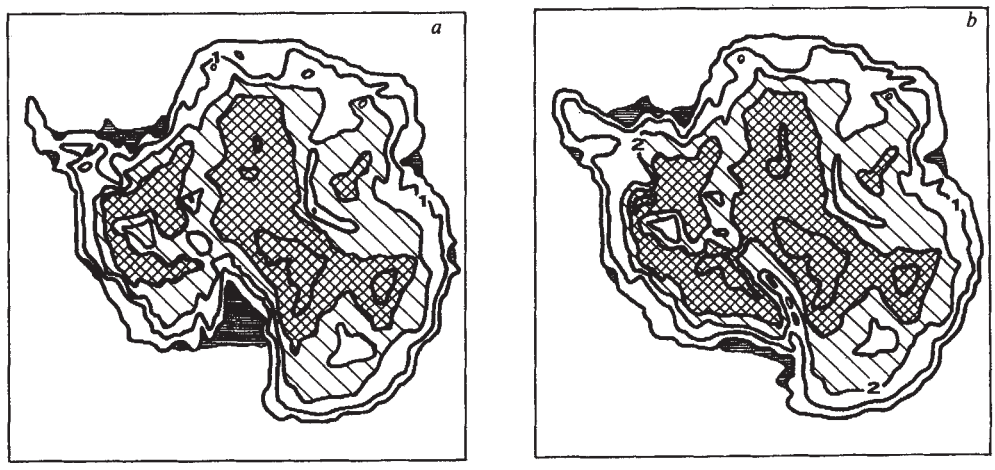


Fig. 3 Equilibrium states of the Antarctic Ice Sheet, computed with the numerical model for present sea level (*a*) and a sea-level stand of 100 m below the present one (*b*). In both integrations the initial condition was $H = 0$. Ice shelves are indicated by horizontal shading. Heavy lines are contours of ice thickness, except the outer one which gives the grounding line. The contour interval is 1 km. Light shading indicates ice thickness > 2 km, heavier shading ice thickness > 3 km.

© 1982 Nature Publishing Group

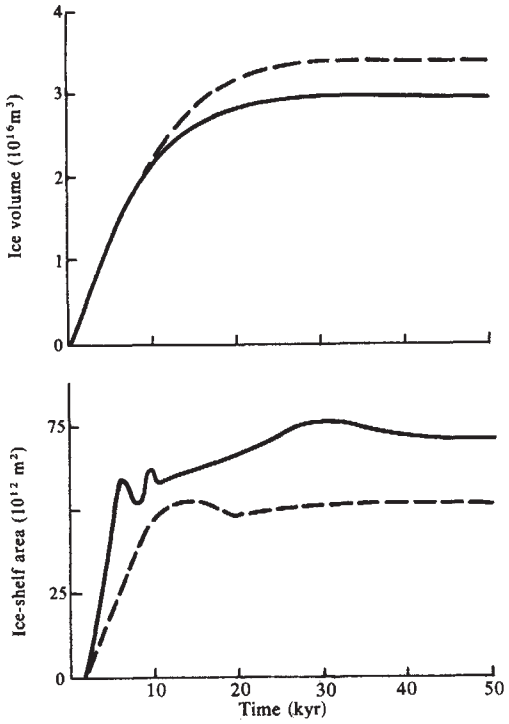


Fig. 4 Total ice volume and area covered by ice shelf, showing how the equilibrium states of Fig. 3 are reached. Solid line, experiment with present sea level; dashed line, experiment with the 100 m drop in sea level.

present distribution of grounded ice appear at the right locations with approximately the right thickness. Although ice-shelf dynamics does not explicitly occur in this procedure, it is dealt with in an implicit way, of course. It is no coincidence that ice shelves occur where the degree of enclosure is large. The compressive stresses associated with large enclosure counteract the ice thinning due to regular spreading. The present procedure mimics this mechanism well.

Because surface temperature and orographic enhancement of precipitation depend on the distribution of ice thickness, a general formulation for the ice accumulation rate G is desirable in which these effects appear in an explicit way. Based on the present distribution of surface elevation, surface slope and accumulation rate[10], the gain of ice mass due to precipitation is written as (in m ice equivalent yr^{-1}):

$$P = \max\,[0.05; f(\theta_0)(0.3 + 14\nabla H' - 4 * 10^{-5} H')/C] \quad (4)$$

So precipitation increases with surface slope and decreases with surface elevation. C is a 'continentality factor', which is computed by checking how many grid points in a circle of radius 500 km are covered by ice or bare ground. C ranges from 1 for an island of one grid point to 2 for the interior of a continent. The function $f(\theta_0)$, where θ_0 is sea-level temperature, models the dependence of snowfall on atmospheric water vapour content and the transition from snow to rain. Its shape is shown in Fig. 2*a*, in which a snowfall curve is drawn for $H' = 0$ and $\nabla H' = 0$. It is based on climatological practice and climatic change experiments with circulation models of the atmosphere[12,13].

Parameterization of ice/snow melt is based on mass-balance studies at the edge of the Greenland Ice Sheet[14]. With the assumption that the yearly cycle in surface temperature is more or less conserved in conditions of climatic change, the following approximate relation holds:

$$\begin{aligned} \text{melt} &= 0 && \text{if } \theta_s < -12\,^\circ\text{C} \\ \text{melt} &= (12 + \theta_s)^2 * 0.028 && \text{if } \theta_s \geqslant -12\,^\circ\text{C} \end{aligned} \quad (5)$$

as for P, the unit is m ice equivalent yr^{-1}; θ_s is surface temperature which can be calculated from θ_0 by using a constant lapse rate γ (so $\theta_s = \theta_0 + \gamma H'$). Equation (5) only holds for temperatures below 0 °C or so, but this is sufficient because higher temperatures do not permit any ice cover to survive.

The difference between snowfall and ice/snow melt makes up the mass balance G. Given the surface slope and elevation, and the continentality factor, it only depends on sea-level temperature. The first three factors are determined by the shape of the ice sheet itself. So climatic change experiments can be performed by varying θ_0 only.

To illustrate how the mass balance depends on local conditions, Fig. 2*b* shows $G(\theta_0)$ for three typical conditions: (I) on the edge of an ice sheet, (II) in the interior of the continent at 2,000 m elevation, (III) in the interior at sea level.

The formulation of the mass balance described above is capable of reproducing all major characteristics of the present mass-balance field over the Antarctic Ice Sheet. It adds considerably to the internal freedom of the model. For example, zones of high precipitation rate follow the edge of the ice sheet, which is certainly realistic.

All model equations are solved with a numerical scheme on a grid with 100-km spacing between the grid points (see Fig. 1*a*). The time step used is 10 yr. Model parameters not yet specified are: $m = 2.5$ and $K = 0.5$ m$^{-3/2}$ yr^{-1} (flow parameters), $T = 4{,}000$ yr (e-folding time scale for bedrock adjustment) and $\gamma = -8$°C km^{-1} (lapse rate). Values of K and m were found by tuning the model to the present Antarctic Ice Sheet for prescribed and fixed bedrock elevation, ice accumulation rate and grounding line. However, the choice is not crucial: other combinations of m and K can be found to give similar results.

Results

In the first experiment the model was run to a steady state starting from no ice cover and bedrock topography h_0. The sea-level temperature was set at −14 °C. It takes ~30,000 yr before a steady state is approached, so the integration was

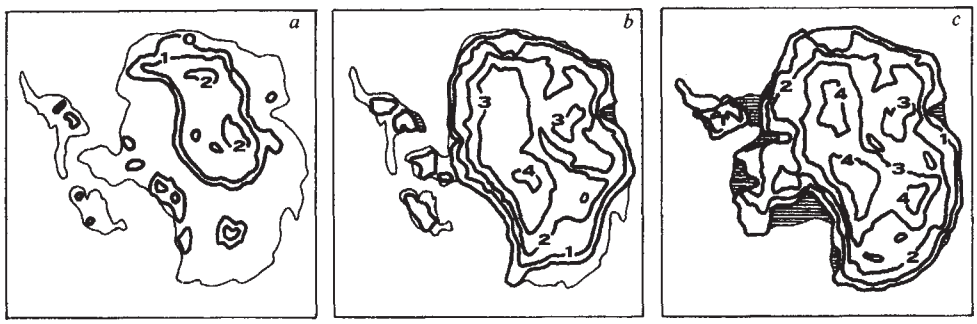


Fig. 5 Equilibrium states of the model Antarctic Ice Sheet for θ_0, the annual sea-level temperature: *a*, $\theta_0 = 5$ °C; *b*, $\theta_0 = 0$ °C; *c*, $\theta_0 = -5$ °C. Heavy lines are contours of ice thickness (km). Ice shelves are indicated by horizontal shading.

© 1982 Nature Publishing Group

carried out over 50,000 yr. This takes about 30 min central processor time on a CDC7600 computer.

The calculated ice distribution is shown in Fig. 3*a*. Comparing this with Fig. 1*a* shows that the model simulation is not too inaccurate, certainly not if one realizes that both bedrock elevations and mass balance are internally generated. Also, note that the state of Fig. 1*a* is not necessarily close to equilibrium, so that a sound judgement of the model performance cannot be made. The major difference between the simulated equilibrium state and the present observed state of the Antarctic Ice Sheet is the extension of grounded ice in the vicinity of the Antarctic Peninsula. In the model simulation a large part of the Weddell Sea is covered by grounded ice. As a consequence, the ice thickness in the western and central parts of the Antarctic Ice Sheet is larger, and the total ice volume is ~25% more than the presently observed volume.

The Ross Ice Shelf is well reproduced by the model, but the slope of the ice surface along the grounding line is steeper than observed. This is probably because the model does not take into account enhanced ice-mass discharge by extensive basal sliding (which is supposed to operate in these regions).

Figure 3*b* also shows the equilibrium state calculated for a sea-level stand which is 100 m below the present one. This was done to investigate the effect of glaciation in the Northern Hemisphere on the Antarctic ice volume. The experiment was identical to the first one (the same initial conditions) except for the 100-m sea-level drop.

Obviously, the West Antarctic Ice Sheet is sensitive to such a sea-level lowering: the grounding line in the Ross Sea advances hundreds of kilometres. At the upper (southern) tip of the Ross Sea, the ice thickness reaches 4 km. Grounding line advance in the Weddell Sea is less spectacular, because in the present-day simulation a large part of the shallow sea is already covered by grounded ice. Altogether, the volume of the West Antarctic Ice Sheet increases dramatically if the sea level lowers by 100 m. According to Fig. 4, the total ice volume increases by ~15%. If sea level is reset to its present value, the ice sheet appears to remain stable. Additional experiments revealed that a sea-level rise of at least 600 m is required to initiate grounding line instability in the Ross Sea (sea level temperature being equal). This suggests that in reality basal sliding must have had an important role in the deglacierization of the Ross Sea at the end of the last glacial extreme, because a 600 m sea-level rise did not occur, of course. Basal melting may have been present during the grounding-line advance, thus keeping the ice thickness of the 'Ross Ice Sheet' small, or it may be triggered by the Holocene climatic warming. However, such considerations are speculative.

So far sea-level temperature θ_0 has been kept constant. Note, however, how the model ice sheet reacts to changes in θ_0. Figure 5 shows steady-state ice sheets computed with $\theta_0 = 5, 0, -5$ °C respectively. For $\theta_0 = 5$ °C, the mass balance is positive only in mountainous regions. In general, the ice does not reach the coast but melts at the ice-sheet edge. For $\theta_0 = 0$ °C, the ice sheet covers most of the eastern part of the Antarctic continent, but ice cover in the western part is still restricted to higher regions. For $\theta_0 = -5$ °C, part of the West Antarctic Ice Sheet is formed, although the (present) Antarctic Peninsula is not yet attached to the main ice sheet. A further drop of θ_0 of a few °C creates the ice sheet shown in Fig. 3*a*. There is not much in between: a steady state in which the Weddell Sea is not covered by grounded ice did not appear.

An interesting feature in these experiments is the large ice thickness occurring for $\theta_0 = -5$ °C in the central parts of East Antarctica (compare with Fig. 3); this is due to the higher snowfall rates (see Fig. 2). However, such findings critically depend on the parameterization of the mass balance and should be considered with caution.

Discussion

The experiments described above should be considered as a first step in developing a more or less complete model of the Antarctic Ice Sheet. Fine tuning of the model has not yet been attempted and several physical processes are not yet included (for example, the impact of extensive basal sliding). So what can we learn from these experiments?

First, it seems that a vertically integrated model is capable of reproducing the influence of mountain ranges on the distribution of ice thickness. This is particularly clear if one considers an experiment in which the bedrock topography is fixed (see ref. 2). In that case the model simulation is very good and errors are only due to the poor resolution of some bedrock irregularities. In fact the value of the flow parameter K used in this study was found by tuning the fixed-bedrock model version to the present distribution of ice thickness. Even if the mass balance and bedrock elevation are free, the simulated ice sheet closely resembles the present Antarctic Ice Sheet. Moreover, because we do not know whether the ice cover in the Weddell Sea region represents an equilibrium state, we cannot state that the model result of Fig. 3 is in error (there is some evidence that ice cover on the Antarctic Peninsula is expanding[15]).

Experiments carried out by Budd and Smith[16] using a similar model of the Antarctic Ice Sheet produced results in agreement with those presented here, although they used a different flow law. This again demonstrates that for large-scale modelling of ice sheets the precise form of the flow law is not important: models can be easily tuned to give similar results.

The Quaternary history of the Antarctic Ice Sheet[17] seems to fit in the picture emerging from the numerical experiments. For a small sea-level drop (due to formation of ice sheets in the Northern Hemisphere) the ice sheet jumps to a state with large ice volume. It returns to a state with small ice volume only if sea level and sea-level temperature are sufficiently high (during a Northern Hemisphere interglacial). However, the required increase in sea level and/or temperature is very large and additional mechanisms are apparently needed to remove the grounded ice from the Ross and Weddell seas. Surging of the West Antarctic Ice Sheet could be one of them.

Future development of this model will include a more refined treatment of ice shelves, computation of the temperature field within the ice and a calculation of the effect of basal sliding on ice-mass discharge. In such a model ice surges may occur. Experiments with a one-dimensional model version including thermodynamics have shown that such events can be modelled very well by vertically integrated models[18]. The computational efficiency of the vertically integrated models (as compared with models with a grid in the vertical direction) will make it possible to study possible surging of the Antarctic Ice Sheet as a 'latitude–longitude problem' instead of looking along a steady flow line[19].

Received 8 February; accepted 20 April 1982.

1. Mahaffy, M. A. W. *J. geophys. Res.* **81**, 1059–1066 (1976).
2. Oerlemans, J. *J. Climatol.* **2**, 1–11 (1982).
3. Birchfield, G. E., Weertman, J. & Lunde, A. T. *Quat. Res.* (in the press).
4. Pollard, D. *Nature* **296**, 334–338 (1982).
5. Weertman, J. *Nature*, **261**, 17–20 (1976).
6. Jenssen, D. *J. Glaciol.* **18**, 373–390 (1977).
7. Budd, W. F., Jenssen, D. & Radok, U. *Derived physical Characteristics of the Antarctic Ice Sheet* (Publ. No. 18 of Department of Meteorology, University of Melbourne, 1970).
8. Nye, J. F. *J. Glaciol.* **3**, 493–507 (1959).
9. Turcotte, D. L. *Adv. Geophys.* 51–86 (1979).
10. *Atlas Antarktiki* (Glawnoje Oeprawlenie geodezii i Kartografii MG, 1966).
11. Greischar, L. L. & Bentley, C. R. *Nature* **283**, 651–654 (1980).
12. Manabe, S. & Stouffer, R. J. *J. geophys. Res.* **85**, 5529–5554 (1980).
13. Oerlemans, J. & Vernekar, A. D. *Contr. atmos. Phys.* **54**, 352–361 (1981).
14. Ambach, W. *Polarforschung* **42**, 18–24 (1972).
15. Sugden, D. E. & Clapperton, C. M. *Nature* **286**, 378–381 (1980).
16. Budd, W. F. & Smith, I. N. *Annals Glaciol.* (in the press).
17. Thomas, R. H. & Bentley, C. R. *Quat. Res.* **10**, 150–170 (1978).
18. Oerlemans, J. *Free Oscillations of Polar Ice Sheets*, Preprint (University of Utrecht, 1982).
19. Budd, W. F. & McInnes, B. J. *Hydr. Sci. Bull.* **24**, 95–104 (1979).

© 1982 Nature Publishing Group

Effects of volatiles on mixing in calc-alkaline magma systems

Herbert E. Huppert*, R. Stephen J. Sparks† & J. Stewart Turner‡

* Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK
† Department of Earth Sciences, University of Cambridge, Downing Site, Cambridge CB2 3EQ, UK
‡ Research School of Earth Sciences, Australian National University, Canberra, Australia

The intimate mixing between different magmas of disparate densities, characteristic of some calc-alkaline magma systems, is explained by a new mechanism involving the emplacement of a layer of wet undersaturated mafic magma at the base of a magma chamber containing more differentiated magma. Turbulent transfer of heat between the mafic magma and the overlying magma leads to crystallization and exsolution of volatiles in the lower layer. For an initial water content of a few per cent the bulk density of the mafic magma can become equal to that of the overlying magma causing overturning and intimate mixing of the magmas. The mechanism is most probable in the low pressure environment of a high level magma chamber.

THE replenishment of a magma chamber by an influx of new magma from depth has been recognized as important in determining petrological and geochemical characteristics of volcanic and plutonic rocks[1] and in triggering volcanic eruptions[2-5]. Previous investigations of the fluid dynamics of replenishment[3,6-9] have been directed mainly towards understanding systems involving dry basic magmas and have not been concerned with effects resulting from the release of gas. In this article we investigate some of the fluid dynamical phenomena due to the presence and release of water vapour which can occur when a magma chamber is replenished by hydrous magma, and consider specifically magmas of the calc-alkaline association. We note that the term 'calc-alkaline' is a common, but imprecisely defined way of describing magmas which are generally associated with orogenic belts and some island arcs. Such magmas are usually, but not ubiquitously, characterized by low amounts of Fe enrichment and relatively high amounts of volatiles.

There is increasing evidence that intimate mixing of magmas is a common phenomenon. In basaltic magmas mixing can be explained in many circumstances because differentiated basaltic magmas can have the same density as higher temperature, more primitive basaltic magmas[3,6,9]. Also the relatively low viscosity of many basaltic magmas allows rapid mixing between them. However, evidence for magma mixing in volcanic rocks of the calc-alkaline association is also common. The belief that mixing has an important role in the evolution of many basaltic andesites, andesites and dacites is based on observations of disequilibrium phenocryst assemblages, inhomogeneous glass compositions, wide ranges of compositions of glass inclusions in neighbouring phenocrysts, banded pumice fragments and mafic clots or micropillows incorporated in more silicic pumice or lava which show distinctive quench textures[5,10-14]. Possibly the most convincing petrographical evidence for intimate mechanical mixing comes from the occurrence of bimodal phenocryst assemblages. Many basaltic to andesitic lavas, for example, contain both calcium-rich and calcium-poor plagioclase phenocrysts and orthopyroxene and augite with high Fe/Mg ratio together with pyroxene with low Fe/Mg ratio[11-14]. Petrological arguments often indicate that mixing between two or more compositionally distinct magma batches must have occurred only shortly before eruption, leading to the suggestion that mixing may even trigger an eruption[2,4,5,10] or alternatively that the mixing was a consequence of the eruption.

If magma mixing is accepted as a widespread process in many calc-alkaline magma systems, there is a major physical problem to solve. Dry magmas of the calc-alkaline association show a steady decrease in density with differentiation along with the decline in temperature[15]. In a conventional model of a volcano, the plumbing system or chamber is filled with magma which gradually differentiates during repose periods. If new undersaturated magma replenishes a chamber from depth (perhaps resulting in an eruption) and forms a lower layer it will, in general, be more primitive, hotter and denser. But the decrease in density of the new magma due to crystallization and thermal effects alone is not sufficient to make the densities of the new and resident magmas equal. Thus there does not at first sight appear to be a driving force in such a situation to cause mixing. This argument, however, assumes that vapour is absent (no free volatiles). Calc-alkaline magmas are frequently rich in water, and Anderson[16] has summarized the evidence which suggests that the water content of many mafic to intermediate calc-alkaline magmas is often in the range 1–4% by weight. Eichelberger[5] has shown that the presence of either dissolved or exsolved volatiles in calc-alkaline magmas has large effects on bulk density. We continue along this line to demonstrate that even a small percentage of volatiles can have profound effects on the fluid dynamics of convection in replenished magma chambers.

We consider in this article the influence of water exsolution on the densities of oversaturated hydrous mafic magmas and show that the induced changes in density can be substantial at low to moderate pressures. This idea follows on from the study of Eichelberger[5] who, influenced by observed density differences of mafic inclusions within rhyolitic lavas, proposed that vesiculation of hydrous mafic magma at the interface between two layers can cause density changes large enough to allow mafic magma to become buoyant in more silicic magma. However, only the relatively thin interfacial region would be destabilized by the process suggested by Eichelberger, and not the entire lower layer. We aim to demonstrate that density changes in the whole layer of saturated hydrous basaltic or andesitic magma can be considerable and that, at low to moderate pressures, intimate mixing can take place between magmas of the calc-alkaline association.

Effects of volatile release on magma density

It follows from basic principles that at equilibrium the bulk density of a hydrous saturated magma in the presence of excess wate is given by

$$\sigma^{-1} = RT(N-n)/P + (1-N+n)/\rho \qquad (1)$$

where σ is the bulk density of the magma, R is the gas constant, T is the temperature in Kelvin, P is the pressure, N is the total weight fraction of water, n is the saturated weight fraction of water dissolved in the melt phase of the magma, and ρ is the

0028-0836/82/240554—04$01.00
© 1982 Macmillan Journals Ltd

weighted density of crystals plus melt. The first term in equation (1) represents the contribution due to the gas and the second term that due to the crystals and melt. Only if $N > n$ is there any gas; otherwise $\sigma = \rho$. The perfect gas law can be used in equation (1) because at the temperatures and pressures under consideration the departures from this law are less than a few per cent.

The weight fraction n can be estimated from data on the saturation concentration of water in silicate melts[17]. The solubility of water can be calculated from the empirical formula

$$n = s(1-x)P^{1/2} \quad (2)$$

where s is the solubility constant and x is the weight fraction of the crystalline phases in the magma.

As an example, we consider the case of hydrous basaltic magma injected into the base of a magma chamber containing more fractionated, cooler and lower density magma (basaltic andesite, andesite or dacite). We make the simplifying assumptions that the crystallization of the basaltic magma involves only anhydrous phases. The bulk density, ρ, of the residual melt and crystals has been calculated by assuming that the crystals have a constant density of 2.8 g cm^{-3} and by evaluating the decrease in density of the residual melt as water is concentrated by crystallization. The density of the melt was calculated using the method of Bottinga and Weill[18]. Variations of ρ with T alone due to thermal expansion are neglected because they are negligibly small in comparison with changes in bulk density of the magma due to volatile exsolution. The solubility constant $s \approx 0.0014$ bar$^{-1/2}$ is taken from solubility data of water in Mount Hood andesite[17].

Finally, we require a relationship between temperature and crystal content. The fractional crystallization of calc-alkaline magmas is quite complex and varies from volcano to volcano. Thus a unique relationship cannot be accurately formulated. However, inspection of estimates on the eruption temperatures and on fractional crystallization models of calc-alkaline magmas suggests that

$$T = 1{,}373 - 200x \quad (3)$$

is a plausible expression[14,19,20]. The results of our density calculations are not in fact strongly dependent on temperature and thus an approximate relationship such as equation (3) seems acceptable.

Figure 1 presents the results of bulk density calculations for a hydrous basalt crystallizing at 0.5, 1.5 and 3.0 kbar pressure with an initial temperature of 1,100 °C and zero initial crystal content. The specific calculations presented are representative of the general behaviour of gas-rich magmas and are similar to the results of Eichelberger[5]. From equation (2), the maximum total weight fraction of dissolved water in a silicate melt is represented by $N = SP^{1/2} \equiv N_0$. For $N < N_0$, the magma is initially undersaturated. It continues to be undersaturated until the weight fraction of crystals $x = 1 - (N/N_0)$, when saturation first occurs. During this stage the bulk density increases slightly as the proportion of crystals in the system increases. Further crystallization involves exsolution of small water vapour bubbles and, assuming these remain suspended in the magma, leads to a continuously decreasing bulk density. In general, if the total water content is decreased, keeping all other variables fixed, the density increases. If the pressure is increased, keeping all other variables fixed, the density of the gas bubbles increases and thus so does the bulk density.

The specific calculations demonstrate that the bulk density of hydrous basaltic magma can vary as a consequence of total water content and substantially as a result of volatile exsolution. The approximate densities of dry degassed magmas of the calc-alkaline association are indicated on the right-hand side of Fig. 1. It follows that at 1.5 kbar, for example, oversaturated basaltic magma can reach the same density as dry, or even volatile undersaturated, andesitic, dacitic or rhyolitic magma. This comparison of hydrous basalt densities with 'dry' densities is merely presented to illustrate the very substantial variations that can occur in the bulk density of a fixed magma composition if excess volatiles are present in the lower layer.

Even at pressures >1.5 kbar the changes can still be large enough to cause density inversions to occur in some magma chambers, although the total water contents required are at the high end of the range reported by Anderson[16]. If water contents between 1 and 4% are typical of calc-alkaline mafic to intermediate magmas, the density changes necessary for convective mixing to take place can occur in shallow-level magma chambers.

Geological applications

When mafic magma is injected into the base of a chamber containing less dense, more differentiated magma, any mixing

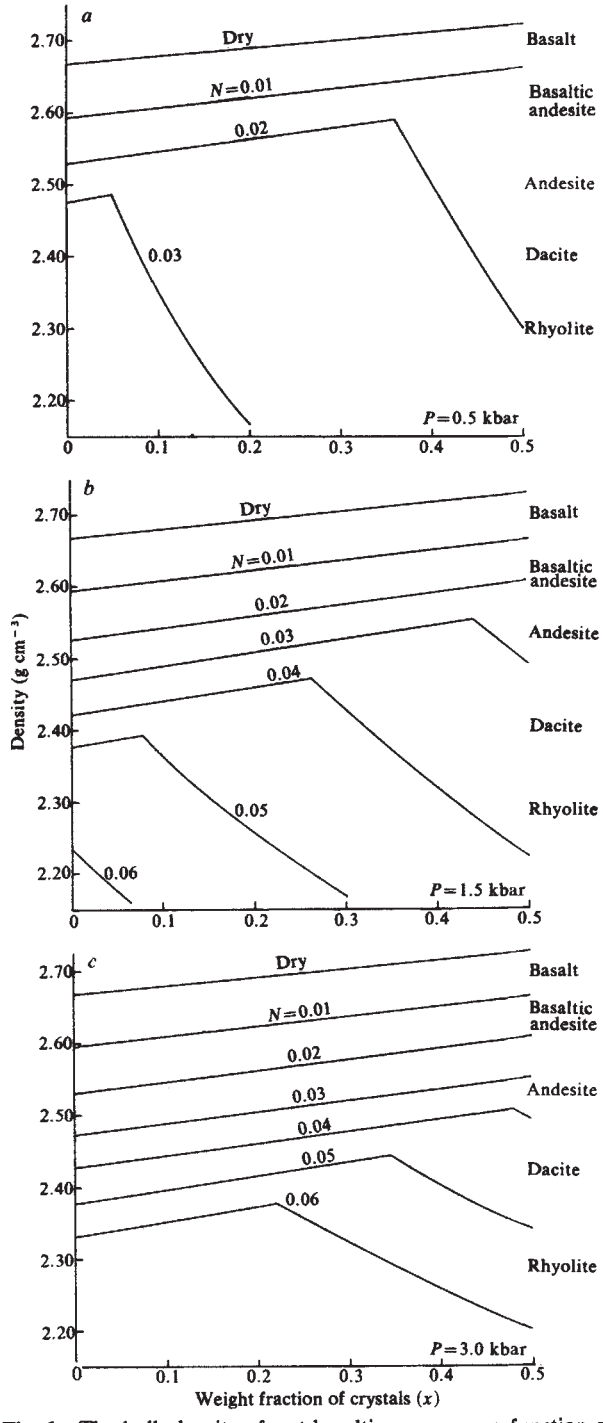


Fig. 1 The bulk density of wet basaltic magma as a function of the weight fraction of crystals for various total weight fractions of water, N, at: *a*, 0.5 kbar; *b*, 1.5 kbar; and *c*, 3.0 kbar.

© 1982 Nature Publishing Group

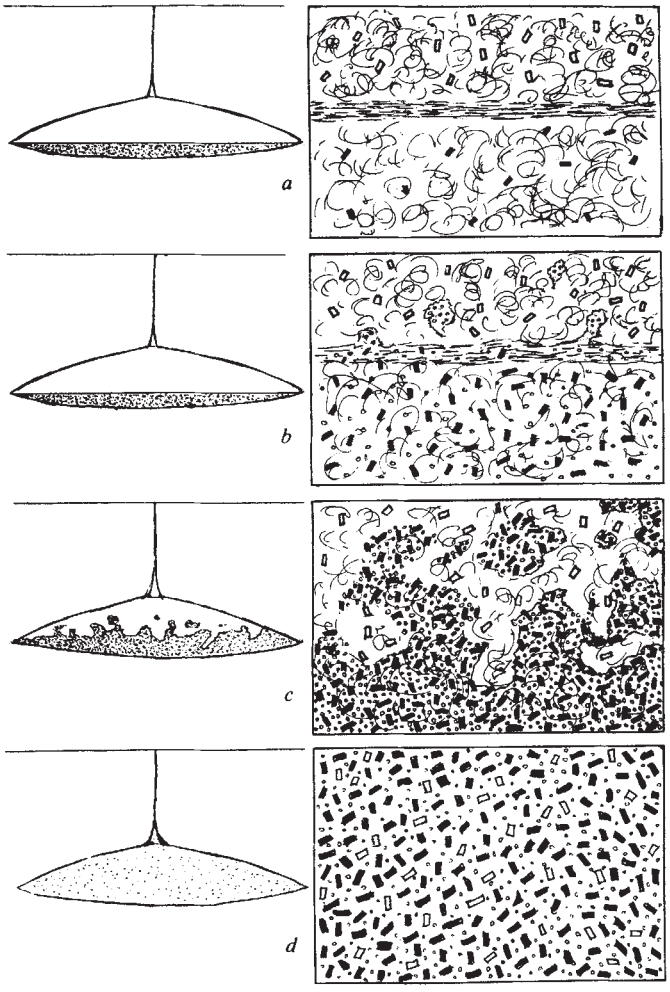


Fig. 2 Schematic model for a magma chamber replenished by an influx of hot, wet basaltic magma. The four diagrams on the right represent close-up views at the height of the interface of the corresponding diagrams on the left. *a*, The heavier magma of the lower layer loses heat across the interface and crystallizes. *b*, The magma of the lower layer becomes saturated and further crystallization produces exsolution of water vapour bubbles. *c*, The bulk density of the magma of the lower layer becomes equal to that of the overlying magma and the lower-layer magma rises into the upper layer. *d*, Intimate mixing between the magmas results.

will initially be strongly inhibited by the density contrast. A two-layered or multi-layered system, with a horizontal interface separating the recently injected hot, denser, mafic magma from the cooler, less dense, silicic magma above, therefore seems a likely precursor to any well-mixed state. Subsequent crystallization is one process for producing a decrease in density of the lower layer. In the case of picritic basalt magmas, crystallization of dense components, notably olivine, leads to a decrease in the density of the residual liquid which can be sufficient to cause overturning and mixing with overlying differentiated basalt[6,7]. In the case of calc-alkaline magmas, changes in the density of the liquid composition due to crystallization alone may not be sufficient to result in overturn. We suggest that it is the release of water vapour once the lower layer has become saturated, which is essential.

The overall process we envisage has a qualitative resemblance to one proposed[6,7] for basaltic magma chambers. We suppose that hydrous mafic magma enters a chamber containing more differentiated magma, as depicted in Fig. 2*a*. This influx may be a consequence of tectonic activity and could possibly trigger an eruption[2,4]. We assume that the volume of new magma forming the lower layer is much less than that of the upper layer and so the evolution of the lower layer proceeds with little change in the upper layer. This is depicted quantitatively for the case of a dry basaltic magma chamber in Fig. 7 of ref. 7. As in our previous analyses of dry basaltic systems, the hydrous mafic magma transfers heat through a sharp interface and crystallizes (Fig. 2*a*). As crystallization proceeds, the weight fraction of dissolved water increases in the residual liquid until the magma is saturated. It can be shown[7,21,22] that the Rayleigh number of layers, even as thin as a few metres, of viscous andesitic magma at a temperature excess of 100 °C often exceeds 10^6. Thus the convective motions will be turbulent, as depicted in Fig. 2.

After the magma saturates, small bubbles form and, along with the crystals, are held in suspension by the vigorous turbulent convective velocities in the lower layer (Fig. 2*b*). The bubbles remain in suspension because the velocity of rise of small gas bubbles relative to quiescent fluid is insignificant in comparison with the r.m.s. turbulent vertical velocities in the lower layer. Similar reasoning applies to the crystals; quantitative calculations for the velocities in this case are given in ref. 7. The exsolution of water vapour causes the bulk density of the lower layer to decrease continually in accord with the results presented in Fig. 1. Eventually, the bulk density may become equal to that in the overlying magma (which has changed by only an insignificant amount), whereupon overturning and hybridization of the magmas occur (Fig. 2*c*, *d*). Further reactions involving the gas phase may be possible during the mixing. We emphasize that we chose the simple and arbitrary chamber geometry depicted in Fig. 2 for schematic purposes only. The processes we have described are fundamental and would also occur in more complex and geometrically irregular plumbing systems. Note that the magma in the upper layer need not be compositionally homogeneous; the same processes would occur if the upper layer were compositionally zoned, though the mixing would be confined to the lower parts of the system[9].

Eichelberger[5] proposes a different mechanism for mixing in which the interface between the mafic and silicic layers vesiculates and becomes unstable. This mechanism seems possible and is incorporated in Fig. 2*b*. However, the interface thickness can be calculated[7] to be quite small, of the order of a few centimetres for temperature differences between layers of ~100 °C, and will thicken to only a few metres as the layers exchange heat. Thus, while instability of a vesiculating interface may account for the occurrence of some mafic inclusions in silicic volcanic rocks, it cannot account for the mixing of large volumes of magma.

There are several further predictions which follow from the general model when applied directly to calc-alkaline systems. For a given pressure the mafic magma must have a water content within a restricted range for the mechanism to be viable. For example, as illustrated in Fig. 1*b*, at 1.5 kbar the total water content for the mafic magma to be initially undersaturated must be $<5.4\%$ and, assuming that a fluid model is applicable only up to a crystal content of 50%, the total water content for saturation to occur must be $>2.8\%$. Overturning into an upper rhyolitic layer of density 2.3 $g\,cm^{-3}$ will not occur unless the total water content exceeds 3.8%. If the upper layer of andesitic magma has degassed, however, the total water content of the mafic magma needs only to exceed 3.0%. Thus there will be situations in the calc-alkaline association where overturning cannot occur and a final compositional stratification will result. Due to the variation of these limits on water content with chamber pressure, with the composition of the more differentiated magma and with the water content of the input mafic magma, both compositional zonation and intimate mixing could occur in the same magmatic system, but at different depths. The data in Fig. 1 suggest that the mechanism we propose is most easily achieved in high-level conduits and near surface chambers beneath volcanoes.

There are considerable viscosity variations amongst calc-alkaline magmas, which will result in large variations in the time scale of the evolution of the layers. A quantitative analysis can be made for the heat transfer in a two-layer system[7] and,

© 1982 Nature Publishing Group

from equation (11a) of ref. 7, the temperature and crystallization history of the lower layer can be predicted. For basaltic systems the calculations suggest that the time taken to overturn can range from weeks to years depending on the initial thickness of the lower layer. In calc-alkaline magmas, silica enrichment can result in increases in viscosity from two to more than four orders of magnitude over the viscosity of tholeiitic basalts. Application of the two-layer system model[6,7] indicates that the time for overturn varies directly as the cube root of the viscosity, and hence overturn times may be increased by factors of order 10 or more for calc-alkaline magmas. Incubation times are then typically years or decades. For more complex layered systems, a similar overturn mechanism can occur[9], but the thermal history of the system cannot yet be quantified. However, overturn times are likely to be considerably longer.

In connection with volcanic activity, we note that in a rigid chamber of fixed volume, exsolution of gas in the lower layer must result in increasing chamber pressure. Blake[4] has shown that an increase of chamber pressure sufficient to exceed the chamber wall fracture strength can occur by influx of new magma of volume typically of the order of 10^{-3} of the chamber volume. Exsolution of a gas phase in the lower layer will result in increasing chamber pressure, and thus the generation of gas due to crystallization could trigger an eruption. For example, if the lower layer constitutes 5% of the total chamber volume and a density decrease of 0.1 g cm^{-3} occurs due to vesiculation in this layer, the volume of gas will be 2×10^{-3} of the chamber volume. Thus many situations can be envisaged where eruption could be triggered by vesiculation in the lower layer.

We also note that the mixing during overturn may produce reactions between the magmas and volatiles which could be significant in driving or triggering volcanic eruptions. Vesiculated mafic magma from the lower layer will be transported during the mixing process to environments which, on average, will be lower in temperature and pressure. Further crystallization of the mafic magma during mixing will yield additional gas release. The bubbles will have to expand, or alternatively increase their internal pressure, which will generate extra chamber pressure to drive magma to the surface.

Most of the convincing petrographical evidence for hybrid magma in the calc-alkaline association comes from stratovolcanoes erupting small-to-medium volumes of mafic to intermediate volcanic products[10,12,14]. In such systems stagnant magma occupies the high-level conduits and reservoir systems beneath the summit during quiescent periods. Such magma will have undergone partial or complete degassing during previous periods of activity and periods of quiescence. We envisage that this magma will become either just saturated in volatiles, or in some cases undersaturated if magma degassed at the surface sinks back into the shallow-level reservoir. Many eruptions from such volcanoes are envisaged to occur during ascent of fresh magma from depth. We suggest that as new magma arrives in the shallow-level system, oversaturation can occur and that mixing will ensue between the old degassed magma and the new volatile-bearing magma.

There are other mechanisms which can produce magma-mixing phenomena. For example, as shown in Fig. 1, a saturated mafic magma containing excess volatiles can be lower in density than the more differentiated resident magma. In this case the new magma is immediately buoyant and ascends as either a laminar or an entraining turbulent plume[3]. Mixing can also occur by sidewall crystallization of a mafic magma forming a turbulent layer of light silicic melt[23,24]. Mixing between the boundary-layer fluid and chamber fluid can form hybrid magma, although there is no evidence that this mechanism can produce large volumes of hybrid magma. There are no doubt other situations which also lead to local mixing within the complex, near-surface plumbing systems of volcanoes. We have tried to document a plausible mechanism which is fluid-dynamically possible, and also consistent with conventional ideas on the plumbing systems of volcanoes. The mechanism is capable of explaining the common occurrence of hybrid calc-alkaline volcanic rocks. In some circumstances basaltic magma can become lighter than even dry rhyolitic magma and thus mixing of mafic magmas derived from the mantle and silicic magmas derived from the crust could occur.

We thank A. T. Anderson, R. J. Arculus, J. C. Eichelberger and B. D. Marsh for helpful comments on a previous draft of the manuscript. The paper is contribution 228 from the Department of Earth Sciences.

Received 9 February; accepted 30 April 1982.

1. O'Hara, M. J. & Matthews, R. E. *J. geol. Soc.* **138**, 237–277 (1981).
2. Sparks, R. S. J. *et al. Nature* **267**, 315–318 (1977).
3. Sparks, R. S. J. *et al. Earth planet. Sci. Lett.* **46**, 419–430 (1980).
4. Blake, S. *Nature* **289**, 783–785 (1981).
5. Eichelberger, J. C. *Nature* **288**, 446–450 (1980).
6. Huppert, H. E. & Sparks, R. S. J. *Nature* **286**, 46–48 (1980).
7. Huppert, H. E. & Sparks, R. S. J. *Contr. Miner. Petrol.* **75**, 279–289 (1980).
8. Huppert, H. E. & Turner, J. S. *Earth planet. Sci. Lett.* **54**, 144–152 (1981).
9. Huppert, H. E. *et al. Earth planet. Sci. Lett.* **57**, 345–357 (1982).
10. Anderson, A. T. *J. Volcan. geotherm. Res.* **1**, 1–33 (1976).
11. Eichelberger, J. C. *Nature* **275**, 21–27 (1978).
12. Luhr, J. F. & Carmichael, I. S. E. *Contr. Miner. Petrol.* **71**, 343–372 (1980).
13. Sakuyama, M. *J. Volcan. geotherm. Res.* **5**, 179–208 (1979).
14. Sakuyama, M. *J. Petrol.* **22**, 553–583 (1981).
15. Bottinga, Y. & Weill, D. F. *Am. J. Sci.* **269**, 169–182 (1970).
16. Anderson, A. T. *J. Geol.* **87**, 509–531 (1979).
17. Burnham, C. W. *Geochim. cosmochim. Acta* **39**, 1077–1084 (1975).
18. Bottinga, Y. & Weill, D. F. *Am. J. Sci.* **269**, 169–182 (1970).
19. Rose, W. I. *et al. J. Geol.* **85**, 63–87 (1977).
20. Arculus, R. J. & Wills, K. J. A. *J. Petrol.* **21**, 743–799 (1980).
21. Bartlett, R. W. *Am. J. Sci.* **267**, 1067–1082 (1969).
22. Rice, A. *J. geophys. Res.* **86**, 405–417 (1981).
23. McBirney, A. R. *J. Volcan. geotherm. Res.* **7**, 357–371 (1980).
24. Turner, J. S. & Gustafson, L. B. *J. Volcan. geotherm. Res.* **11**, 93–125 (1981).

Population dynamics of human helminth infections: control by chemotherapy

Roy M. Anderson* & Robert M. May†

* Department of Pure and Applied Biology, Imperial College, London University, London SW7 2BB, UK
† Department of Biology, Princeton University, Princeton, New Jersey 08544, USA

An analysis is presented of the population dynamics of the major helminth parasites of man with the aim of understanding observed patterns in the age-specific prevalence and intensity of infection. Mathematical models are used to investigate the possibility of controlling helminth diseases by mass chemotherapy, and to explore the advantages of selective treatment of the most heavily infected individuals in a community.

OVER 30 years ago Stoll[1] noted that the major helminth infections of man were among the most prevalent of all human infectious diseases. Today the global picture is similar, and we still live in a very 'wormy world': there are roughly 1,000 million cases each of roundworm (*Ascaris lumbricoides*), pinworm (*Enterobius vermicularis*) and whipworm (*Trichuris trichuria*); over 600 million cases of hookworm (*Necator americanus* and *Ancylostoma duodenale*); 300 million cases of filarial infections

0028–0836/82/240557—07$01.00

© 1982 Macmillan Journals Ltd

(principally *Wuchereria bancrofti* and *Onchocerca volvulus*) and over 200 million cases (500 million people at risk) of schistosomiasis (*Schistosoma mansoni, Schistosoma haematobium, Schistosoma japonicum* and *Schistosoma intercalatum*)[2-5].

In contrast to most viral and bacterial diseases, helminth populations exhibit a remarkable degree of temporal stability in human communities[6,7]. These parasite populations tend to be robust to perturbation, resulting from either seasonal changes in climate or the introduction of control measures. This is in part a consequence of the human host's inability to acquire strong immunity to re-infection, so that helminth infections are persistent in character with hosts being continually re-infected. In endemic regions of the world, a newborn child can expect to harbour worms for most of its life (Fig. 1).

The wide range of methods used in attempts to control helminth infections include chemotherapy, improved sanitation and hygiene, education and (in the case of indirectly transmitted infections such as schistosomiasis and filariasis) vector control[8-10]. New techniques, based on the use of monoclonal antibodies, raise the exciting possibility that vaccination against human helminths may be feasible in the future[11]. Today, safe, effective and cheap chemotherapeutic agents are available for many of the major helminth infections of man (excluding the filarial worms)[12,13]. There remains, however, a need for a better understanding of the most cost-effective ways of delivering these drugs, not at the individual level, but in the community as a whole. Among the unresolved epidemiological questions concerning the use of antihelmintic agents are the following. What proportion of the population must be treated, and how often, to achieve a defined reduction in parasite prevalence and abundance? How rapidly will the parasite population return to its pre-control level once chemotherapy ceases? Is the selective treatment of the more heavily infected individuals of greater benefit to the community, both in terms of costs and of reducing the prevalence of disease symptoms, when compared with random or blanket chemotherapy? This article examines certain theoretical and observational aspects of these questions.

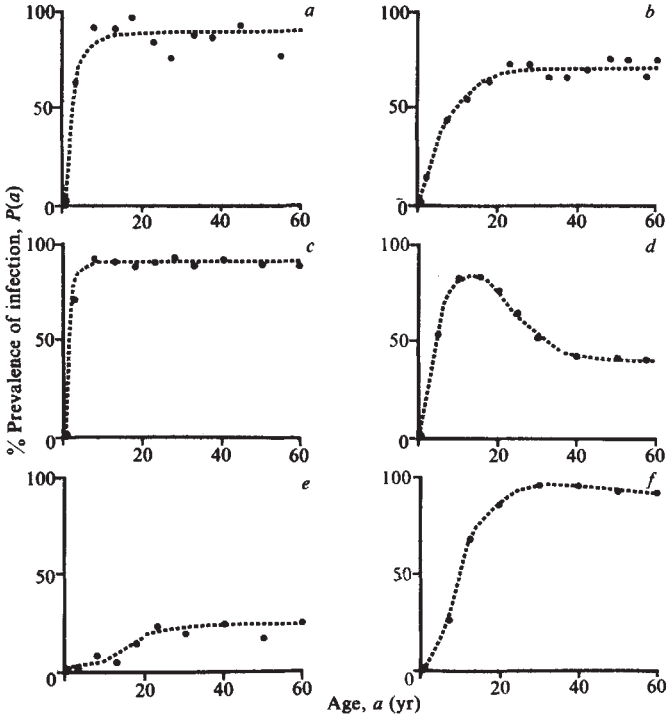


Fig. 1 Examples of horizontal age–prevalence data for some major helminth parasites of man in communities where the infections are endemic: *a*, *A. lumbricoides* in Iran[16]; *b*, hookworm (both *Necator* and *Ancylostoma*) in Taiwan[60]; *c*, *T. trichuria* in the Philippines[61]; *d*, *S. haematobium* in Tanzania[34]; *e*, *W. bancrofti* in India[62] (prevalence of microfilaria carriers); *f*, *O. volvulus* in Mali[49]. The dashed lines are fitted by eye to illustrate the general trends; the solid dots are observed values.

Population dynamics

An understanding of the factors which regulate parasite abundance, and contribute to the observed stability of helminth populations, is of obvious significance to the design of optimal control policies[14,15]. Recent attempts to extend the pioneering work of Macdonald can be summarized as identifying five principal factors controlling the observed dynamics of the major helminth infections of man[7,16-18].

(1) The relative time scales on which the dynamics of the host and parasite populations operate. These are largely determined by the expected life spans of the human host and the various developmental stages of the parasite. The human host typically has a life span an order of magnitude or more in excess of any of the parasitic stages (Table 1*a*). For example, the adult worms in the human host have life spans of 1–16 yr depending on the species of parasite. Note, however, that measurement of this parameter is difficult in practice and hence the values in Table 1*a* are only approximations. Free-living infective stages plus infected vectors have much shorter life expectancies than those of the adult parasites. These empirical observations enable certain simplifications to be made in the analysis of parasite population behaviour, without any serious loss of detail or accuracy[7,15-17]. Given the information in Table 1*a*, we can examine the dynamics of the adult parasites under the assumptions that the host population is effectively constant in size on a time scale appropriate to changes in adult parasite population size and that the populations of the infective stages (or vectors) are essentially at equilibrium due to the rapidity with which changes in these populations occur when compared with those in the adult worm populations[14,15].

(2) The regulation of parasite population growth in the human host. Although man is unable to develop an effective protective immunity to re-infection, he can mount immunological and nonspecific responses to the invasion of helminth parasites, which act in conjunction with resource limitation to reduce worm survival, establishment and fecundity in a manner that depends on the density of parasites in the host (refs 19, 20; Fig. 2). These density-dependent responses are the major regulatory mechanisms controlling the growth of the total parasite population in any human community. The net severity of these constraints in the host population depends critically on the statistical distribution of worm numbers per host[21-23].

(3) Worm distributions: these are typically highly aggregated[21,22,24], with the majority of worms being harboured by a minority of the human population. The generative mechanisms of such patterns are many and varied but include genetic, spatial and behavioural factors. The negative binomial probability distribution has proved to be a good empirical model for observed patterns; this distribution is defined by two parameters, the mean worm burden per host, M, and a parameter, k, which varies inversely with the degree of parasite aggregation. Observed k values are very small, and it is not unusual for >65% of the worm population to be harboured by <15% of the human community (Table 1*b*). This high degree of contagion enhances the severity of density-dependent constraints on parasite population growth[23].

(4) The reproductive biology of the parasite. The major human helminths are dioecious and the production of transmission stages therefore depends on the probability that a female worm is mated. This probability is a function of the statistical distribution of parasite numbers per host; there exists a critical average worm burden, termed the breakpoint, below which mating frequency is too low to maintain the parasite[7,17,25,26]. The precise level of this breakpoint is in part determined by the degree of worm contagion in the human population. Recent studies of *Ascaris* and hookworm infections[7,16], both of whose distributions tend to be highly aggregated in human populations, suggest that the breakpoint is typically very low—of the order of 0.3–0.5 worms per host. Mating prob-

© 1982 Nature Publishing Group

abilities thus seem to be of limited significance to the overall population stability of helminth parasites, and the breakpoint concept therefore appears to have little practical significance for disease control[7,16,26].

(5) The reproductive or transmission potential of the parasite. The basic reproductive rate, R_0, is used to measure this potential and is defined as the average number of female offspring produced throughout the lifetime of a mature female parasite which themselves achieve reproductive maturity, in the absence of density-dependent constraints[7,16,25,27]. If the parasite population reaches a steady state (stable endemic disease) the effective reproductive rate, R, is equal to unity; in other words, each female parasite exactly replaces herself in the next generation. R_0 has recently been estimated for several helminth infections in different regions of the world[7,16,28]; these estimates typically lie in the range 1–5 and are generally lower than those recorded for many viral, bacterial and protozoan infections of man[29,30,70] (Table 1*c*). The condition $R_0 = 1$ defines a transmission threshold below which the parasite is unable to maintain itself in the human population[7,16,25]. The value of R_0 is determined by many individual components representing the birth, death and transmission rates of the various developmental stages in the parasites' life cycle. Estimates of R_0, however, can be obtained directly from epidemiological data recording the prevalence and intensity of infection in different age classes of the host population. The parameter R_0, in conjunction with the life expectancy of the adult parasite, A, the degree of worm aggregation, k, and the severity of density-dependent constraints on worm population growth, z, acts to determine the average intensity, M^*, and prevalence, P^*, of infection at equilibrium in a given community.

Mathematical models

It follows from the relative time scales recorded in Table 1*a* that a description of the temporal dynamics of parasite population growth can be given in terms of one variable, $M(t)$, the mean adult worm abundance per host at time t[7,14,15]. Assuming the parasites to be distributed in a negative binomial manner in the human population, with clumping parameter, k, the dynamics of infection by helminths in any defined human community may be described by a single differential equation for $M(t)$, of the general form

$$dM/dt = (M/A)\,[R_0 f(M,k,z) - 1] \qquad (1)$$

R_0, A and k are as defined above, z is an inverse measure of the severity of density dependence (see Fig. 2) while f denotes the net force of these constraints acting on parasite population growth in an individual host. For example, in ascariasis, density-dependent processes seem to act mainly on worm fecundity[16], such that, within a given host, the rate of egg production per female worm decays approximately exponentially as worm burden increases (Fig. 2). This gives

$$f(M,k,z) = [1 + (1-z)M/k]^{-(k+1)} \qquad (2)$$

and the equilibrium average worm burden, M^*, and prevalence, P^*, are given by

$$M^* = [R_0^{1/(k+1)} - 1]\,[k/(1-z)] \qquad (3)$$

and

$$P^* = 1 - [1 + M^*/k]^{-k} \qquad (4)$$

If density-dependent constraints severely restrict worm fecundity the hosts with the maximum egg output may not

Table 1 Characteristics of the population biologies of human helminth parasites

a, Time scales

Parasite	Life expectancy: Man (yr)	Life expectancy: Adult parasite in man (yr)	Life expectancy: Infective stage or intermediate host	Ref.
A. lumbricoides	40–60	1–2	1–3 months (egg)	16
T. trichuria	40–60	2–3	1–2 months (egg)	8
N. americanus	40–60	2–4	3–10 days (L_3 larvae)	7
S. mansoni	40–60	3–4	3–6 weeks (infected snail)	55, 56
O. volvulus	40–60	8–14	1–3 weeks (arthropod vector)	8, 28, 57
W. bancrofti	40–60	12–16	1–2 weeks (arthropod vector)	8, 58

b, Degree of aggregation of adult parasites in the human population, k

Parasite	Geographical location	k*	Ref.
A. lumbricoides	Iran	0.2–2.9	16
N. americanus	India	0.03–0.6	7
N. americanus and *A. duodenale*	Taiwan	0.05–0.4	7
S. mansoni	Brazil	0.03–0.5	59

c, Basic reproductive rate, R_0

Parasite	Geographical location	R_0	Ref.
O. volvulus	Sudan	5–35	28
A. lumbricoides	Iran	4–5	16
N. americanus	India	2–3	7
S. mansoni	Brazil	1–2	59

* (k, An inverse measure of aggregation.)

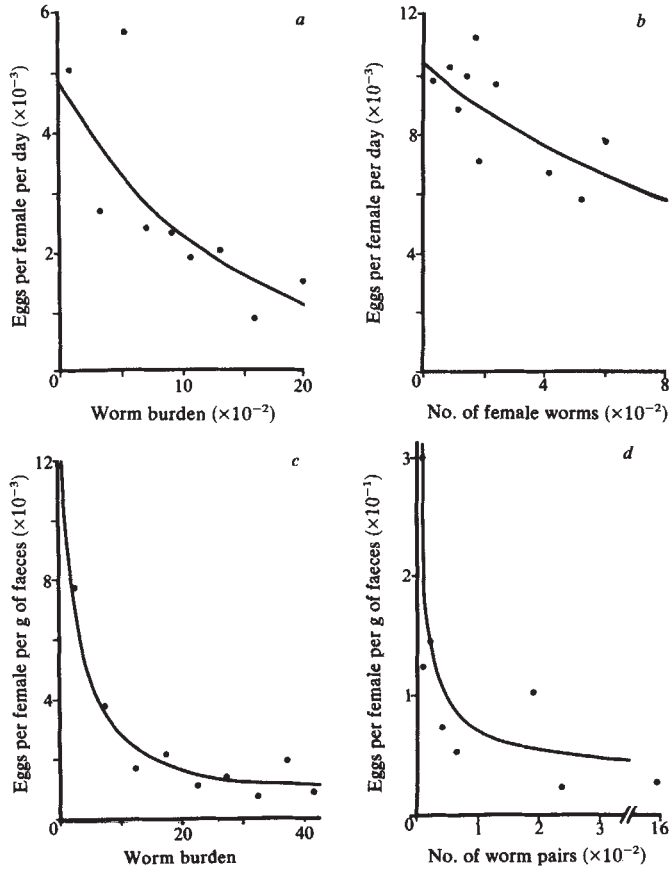


Fig. 2 Examples of the decline in worm fecundity (in terms of egg output) as the burden of parasites in the human host increases (density-dependent egg production). ●, Observed values; —, best-fit power or exponential models. *a*, *N. americanus* (data from ref. 63). The solid line denotes the best-fit exponential model of the form $y = \hat{a}\,e^{-\hat{b}i}$ where y is the egg output per female worm per day, i is worm burden per host, and $\hat{a}$ and $\hat{b}$ are constants with values $\hat{a} = 4.8\times10^3$, $\hat{b} = 0.73\times10^{-3}$ ($r^2 = 0.7$). *b*, Mixed infections of *N. americanus* and *A. duodenale* (data from ref. 64). The solid line denotes the best-fit exponential model, as defined for *a*, with parameter values $\hat{a} = 10.3\times10^3$, $\hat{b} = 0.7\times10^{-3}$ ($r^2 = 0.7$). *c*, *A. lumbricoides* (data from ref. 16). The solid line denotes the best-fit power model of the form $y = \bar{a}i^{\bar{b}}$ where y and i are as defined for *a*, and $\bar{a}$ and $\bar{b}$ are constants with values $\bar{a} = 9.0\times10^{-3}$, $\bar{b} = -0.6$ ($r^2 = 0.9$). *d*, *S. mansoni* (data from refs 59, 65). The solid lines denote the best-fit power model as defined for *c* with parameter values $\bar{a} = 38.4$, $\bar{b} = -0.36$ ($r^2 = 0.7$).

© 1982 Nature Publishing Group

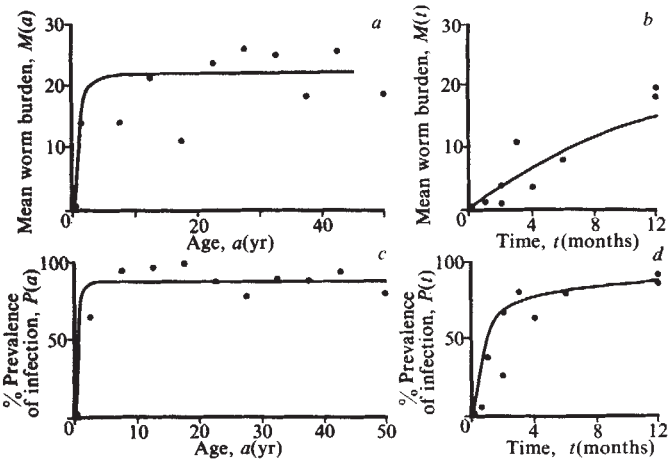


Fig. 3 The fit of the model defined by equations (8) and (9) in the text to data of the epidemiology of *A. lumbricoides* in Iran (data from ref. 16). ●, Observed values; —, model predictions. *a, c*, Rise in the average worm burden and prevalence of infection, respectively, with host age. The data were collected by means of a horizontal epidemiological survey[16]. *b, d* Show, respectively, the increase in parasite intensity and prevalence with time in a sample of people who were cleared of worms by antihelmintic treatment at time $t=0$. Note the rapidity with which both the average worm burden and prevalence approach their pre-control levels (indicated by the plateau levels in *a, c*). Estimates of the parameters of the model [equations (8), (9)] are as follows: $R_0=4.3$, $M^*=22$, $A=1$ y, $z=0.96$, $k=0.57$. The density-dependent function, *f*, used in the model was as defined in equation (2) and the parameter values listed above were used to make the predictions in all four graphs, *a–d*.

necessarily harbour the maximum parasite burden. However, given the typical maximum worm burdens recorded in human communities, the observed patterns of density-dependent fecundity (Fig. 2) suggest that such situations are unlikely to arise in practice. The qualitative dynamical behaviour of the model defined in equation (1) is independent of the precise form of function *f*. The quantitative properties, however, will depend on whether density-dependent processes act on worm fecundity, survival or establishment, or any combination of these. Human helminth fecundity is invariably density dependent (see Fig. 2) but information on survival and establishment is extremely limited, and practically difficult to obtain.

Equation (1) may be used to approximate the dynamics not only of directly transmitted infections (such as *Ascaris*), but also of indirectly transmitted ones (such as *Schistosoma*) since the parameter R_0 represents the appropriate combination of transmission factors for the complete life cycle of any given parasite species[7,14,15,28,31]. The epidemiological data most commonly available are from horizontal and longitudinal studies of age–prevalence and age–intensity relationships[32]. The model can be extended to describe the way the mean worm burden, $M(t,a)$, changes with time, (longitudinal changes), and host age, a (horizontal changes)[28]:

$$\partial M/\partial t+\partial M/\partial a=\Lambda(t)-M/A \qquad (5)$$

The function $\Lambda(t)$ is defined for notational convenience as

$$\Lambda(t)=(R_0/AL)\int_0^\infty l(a)\,M(t,a)\,f(M,k,z)\,\mathrm{d}a \qquad (6)$$

where $l(a)$ is the probability that a host survives to age a and L denotes human life expectancy ($L=\int_0^\infty l(a)\,\mathrm{d}a$). For human communities in which the prevalence and average intensity of infection have remained relatively constant with time (where $\partial M/\partial t\simeq 0$), equation (5) reduces to a simple differential equation for the change in average worm burden, $M(a)$, with host age; in this equation, Λ is a constant defined by equation (6) with $M(t,a)$ replaced by $M(a)$. For endemic infections of *Ascaris* and hookworm, the average worm burden rises rapidly as host age increases to reach a plateau such that the mean worm burden, $M(a)$, in the majority of age classes is essentially equal to the mean worm burden of the total population [the M^* of equation (3)]; consequently, a good estimate of Λ is given by

$$\Lambda=(R_0M^*/A)\,f(M^*,k,z) \qquad (7)$$

It is here assumed that the distribution of parasites in the entire community (consisting of all age classes) remains approximately negative binomial in form with mean M^* and clumping parameter k. The precise overall distribution will generally be formed from a mixture of negative binomial distributions (one for each age class) as a consequence of age-dependent human mortality and differences in exposure to infection[33]. The use of a common aggregation parameter for the entire community, however, remains a good approximation provided parasite life expectancy is short in relation to that of man[33]. Available empirical evidence suggests that the negative binomial distribution remains a reasonable empirical model of hookworm and roundworm numbers per host even when a sample consists of a mixture of individuals from different host age classes[7,16,34].

In this limiting case of steady infection, when $\partial M(t,0)/\partial t\simeq 0$, the appropriately simplified form of equation (5) yields an age–intensity relationship of the form

$$M(a)=\Lambda A\,[1-e^{-a/A}] \qquad (8)$$

Here Λ is given by equation (7). The predicted age–prevalence curve (the proportion of each age class infected) then follows as

$$P(a)=1-[1+M(a)/k]^{-k} \qquad (9)$$

Equations (8) and (9) can be used to obtain estimates of Λ, A and R [using equation (7)] from observed age–prevalence and age–intensity data[7,16,28,35].

We have assumed that the net rate of infection, Λ, is independent of host age and depends only on the average number of transmission stages present in the host's environments. This is often the case for infections such as *Ascaris*[14], but in many instances, Λ varies with age as a result of changes in the contact rate with infective stages (or infected vectors), the severity of density-dependent factors (mediated by age-related changes in host resistance to infection) or the degree of worm contagion in different age classes of the population[7,28,31]. For schistosome and filarial infections, age-related changes in acquired resistance and contact with infection are known to be important, although much controversy still surrounds their relative significance[28,36] (Fig. 1).

The model defined in equations (8) and (9) can provide an accurate description of observed trends[16]. Figure 3 presents data pertaining to the epidemiology of *Ascaris* in Iran. The model can be seen to mimic changes both in the observed horizontal trends of intensity and prevalence in different age classes of the population (Fig. 3*a, b*) and in the re-infection with time of individuals who were given a single treatment with a chemotherapeutic agent.

Control by chemotherapy

Effective and safe antihelmintic agents are available for the treatment of most major helminth infections of man[12,37,38], including mebendazole[5,39,40,66,67] (for enterobiasis, ascariasis and trichuriasis), pyrantel pamoate[8] (for ascariasis, enterobiasis, trichuriasis and hookworm), levamisole[40,41] (for ascariasis), bephenium hydroxynaphthoate[8] (for *Ancylostoma*) and the recently developed praziquantel[42–45] (for schistosomiasis). However, questions concerning the best way to use these drugs for the benefit of the community rather than the individual remain to be answered[10]. The above models can be adapted to examine these issues: we focus first on the long-term effects of the use of mass chemotherapy and, second, on the selective treatment of the most heavily infected individuals in a community.

Mass chemotherapy in the long term. When a drug is administered randomly within a population by treating a pro-

© 1982 Nature Publishing Group

portion, g, of the community per unit of time and the drug has an efficacy h (where h represents the average proportion of the worm burden killed by a single, or short course of treatment), the resulting increase in the per capita death rate of adult parasitic worms, c, is then[7,16]

$$c = -\ln(1 - gh) \tag{10}$$

By subtracting the net death rate due to treatment, cM, from the right-hand side of equation (1), the average worm burden, $\bar{M}$, at the new equilibrium established by the control programme can be shown to be

$$\bar{M} = k([R_0/[1 - A\ln(1 - gh)]]^{1/(k+1)} - 1)/(1 - z) \tag{11}$$

It is assumed that density-dependent constraints act in the manner defined in equation (2); other assumptions give similar results. To eradicate the infection the effective reproductive rate R (the basic reproductive rate, R_0, modified by the action of chemotherapy) must be reduced below unity. To achieve this, the proportion treated per unit of time must be greater than a critical value, $\hat{g}$ (ref. 7):

$$\hat{g} = (1 - \exp[(1 - R_0)/A])/h \tag{12}$$

Even when $g > \hat{g}$ treatment must be continued for a period of time greater than the maximum life span of the adult worm. If treatment ceases before the average worm burden is below the breakpoint (which, unfortunately, we have seen is usually close to zero worms per host), the parasite population will rapidly return to its pre-control level (the speed of return depends on the values of R_0 and A). For helminths, however, eradication is not always a practical goal, and control policies are more often oriented to the elimination of disease rather than of infection because the two are not necessarily equivalent[3,5,45,46,68].

The model defined in equation (5) can be used to examine the impact of a defined programme of chemotherapy on the prevalence and intensity of infection both with time and in different age classes of the population (Fig. 4). The numerical results shown in Fig. 4 mimic the dynamics of endemic *Ascaris* and illustrate two important points—the rapidity with which certain helminth infections return to their pre-control level once chemotherapy ceases[47], and the differential impact of antihelmintic treatment on the prevalence and intensity of infection. Simply monitoring prevalence may lead to the incorect conclusion that control measures have little impact on parasite abundance. The principal advantage of these models is that (given estimates of parameters R_0, A, k and z) they can be used to predict quantitatively the outcome of a defined programme of chemotherapy and to define the level of treatment required for eradication.

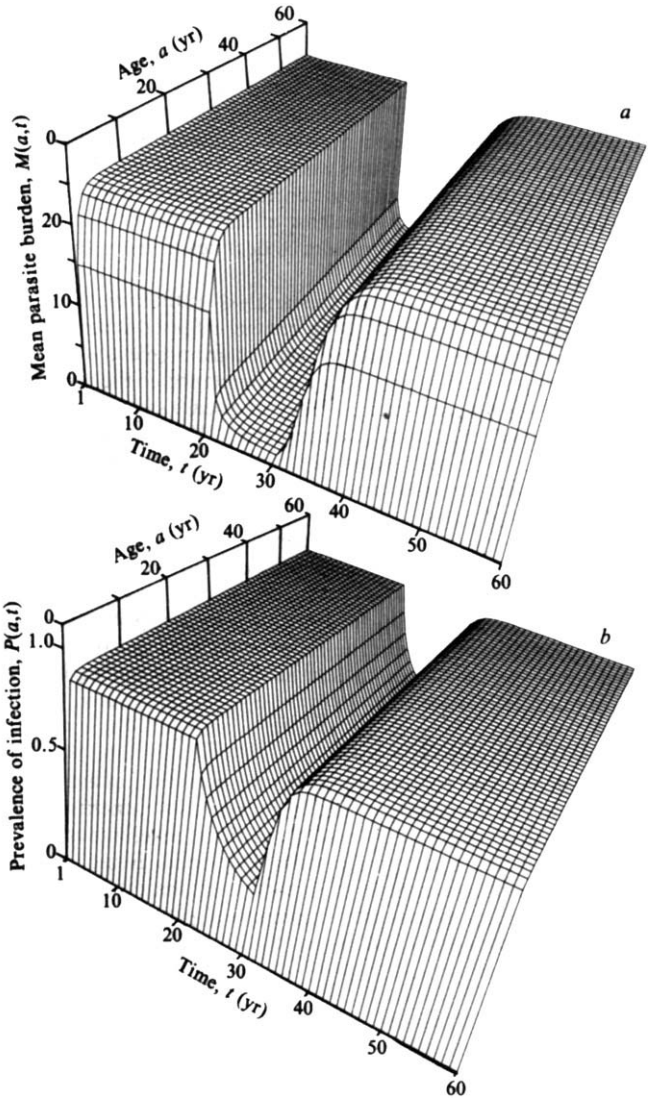


Fig. 4 Two three-dimensional graphs obtained by the numerical solution of the partial differential equation defined in equation (5). The simulations are designed to mimic the dynamics of endemic ascarasis in a human community, both with time and in different age classes of the population. *a*, Changes in the average worm burden; *b*, changes in prevalence. For the first 10 yr the helminth infection is at its endemic equilibrium (with M^* in the older age classes set at 22 worms per host). In year 10, mass chemotherapy is initiated and applied randomly in the community on a continual basis where 20% of the population is treated each month ($g = 0.25$ per month) with a drug of 95% efficacy ($h = 0.95$). After 10 yr the treatment is stopped and the graphs display the return of the parasite population to its pre-control equilibrium level. The human survivorship function, $l(a)$, was chosen to mimic typical patterns in the developing world with a life expectancy, L, of 50 yr. The density-dependent function f was as defined in equation (2) and the remaining parameters of the model were set at the following values: $R_0 = 4.3$, $k = 0.57$, $z = 0.96$, $A = 1$ yr. Note the rapidity with which the infection returns to its pre-control level once chemotherapy ceases and the differential impact of control on the average worm burden, $M(t,a)$, and the prevalence of infection, $P(t,a)$.

Single applications of mass chemotherapy. Antihelmintic drugs are often used intensively over short periods of time to reduce the prevalence of disease symptoms in a community. Frequently, such applications are given at irregular intervals with long time periods between successive treatments. The objectives are to maximize the short-term benefits and minimize the number of treatments (and hence cost). In 1924, Smillie[48] argued that the optimum way to apply mass chemotherapy was selectively to treat the most heavily infected individuals; Warren has recently re-opened this debate[5,49–51,69]. A major disadvantage of this strategy can be the extra cost involved in identifying the most heavily infected individuals; in practical terms this can be done either by faecal egg, urine egg or blood larval counts, or by noting the occurrence of clinical symptoms of disease. Conversely, there are advantages in targeted treatment if the drug has side effects, or if there are any worries about the evolution of drug resistant strains of the parasite[52].

We can use our models to calculate the proportional reduction achieved by mass chemotherapy, in quantities such as the average worm burden, the prevalence and the average egg output. We denote the levels of these parameters before control as M^*, P^* and E^* respectively and define the probability that an individual harbours i parasites as $p(i)$, and that he receives treatment as $g(i)$, where

$$g(i) = a[1 - (1 - \alpha)\,e^{-i/I}] \tag{13}$$

a and α (both < 1) are constants defining the upper and lower bounds on $g(i)$, respectively, while I is a measure of the selectivity of the treatment. For example if $I = 40$, ~60% of individuals with 40 worms receive drug treatment. A continuous function for $g(i)$ is more realistic than a step function, as in

© 1982 Nature Publishing Group

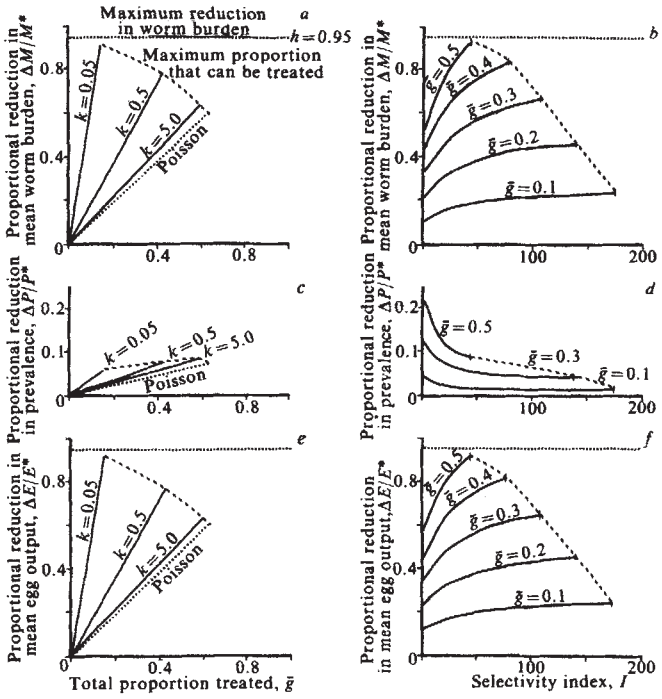


Fig. 5 The impact of selective chemotherapy on the proportional reduction in mean worm burden, $\Delta M/M^*$, the prevalence of infection, $\Delta P/P^*$, and the average egg output per host, $\Delta E/E^*$. *a*, *b*, Relationships between the proportional reduction in average worm burden immediately after a single mass treatment in a community in relation to *a*, the proportion of the population treated, $\bar{g}$, and *b*, the degree of selectivity in treatment, *I*, where the selective function is as defined by equation (13), for various assumptions about the distribution of worms in the human population (that is, various values of the negative binomial clumping parameter, *k*). The relationships in *a* and *b* were calculated using equation (14) with the following parameter values: $M^*=40$, $I=40$, $\alpha=0$, $a=1$, $h=0.95$ and $k=0.5$ (in *b*). *c* and *d*, *f* and *e* show, respectively, similar relationships for the proportional reductions in prevalence and average egg output. The proportional reduction in prevalence, $\Delta P/P^*$ is defined by the relationships $\Delta P/P^* = aJ(I, M^*, k)/[1-(1+(M^*/k))^{-k}]$ where $J(I, M^*, k)=[1+(1-h)(M^*/k)]^{-k}-(1-\alpha)[1+(1-y)(M^*/k)]^{-k}-\alpha(1+(M^*/k))^{-k}]$ and $y=h\exp(-1/I)$. The parameter values for *d* and *c* are as defined for *a* and *b*. The proportional reduction in average egg output, $\Delta E/E^*$, is defined by the relationship $\Delta E/E^* = aK(I, M^*, k)/D(I, M^*, k)$ where $K(I, M^*, k)=[[1+(1-x)(M^*/k)]^{-(k+1)}-(1-\alpha)y[1+(1-xy)(M^*/k)]^{-(k+1)}]$, $D(I, M^*, k)=[1+(1-x)(M^*/k)]^{-(k+1)}$ and $x=\exp(-\bar{b})$ where $\bar{b}$ is as defined in Fig. 2 legend and *y* is as defined above. The parameter values for *e* and *f* are as defined for *a* and *b* excepting *h* is set at unity.

practice there is always considerable uncertainty about the worm burden harboured by any given individual. The proportional reduction in the average worm burden, $\Delta M/M^*$ [where $\Delta M = h\sum_{i=0}^{\infty} ip(i)g(i)$], achieved by treating a proportion $\bar{g}$ [where $\bar{g}=\sum_{i=0}^{\infty} g(i)p(i)$] of the total population with selectivity *I* is given by

$$\Delta M/M^* = \bar{g}\,h\frac{G(I, M^*, k)}{H(I, M^*, k)} \qquad (14)$$

where *h*, M^* and *k* are as defined earlier. The functions *G* and *H* are defined as

$$G(I, M^*, k)=[1-\hat{z}(1-\alpha)[1+(1-\hat{z})(M^*/k)]^{-(k+1)}] \qquad (15)$$

and

$$H(I, M^*, k)=[1-(1-\alpha)[1+(1-\hat{z})(M^*/k)]^{-k}] \qquad (16)$$

where $\hat{z}=\exp(-1/I)$.

Figure 5 displays the impact of various programmes, using different selectivity levels (*I*), and various proportions treated ($\bar{g}$) on the average worm burden, the prevalence and the average egg output, given various assumptions concerning the degree of worm clumping in the human population. Three points emerge from this analysis.

(1) Selective treatment is highly beneficial (in terms of the number of treatments administered), provided the worms are highly clumped in their distribution in the host population (in practice this is indeed the case; see Table 2*b*). For example, in the case shown in Fig. 5*a*, where the selectivity index is set at 40 worms per host, a 50% reduction in average worm burden ($M^*=40$) is achieved by treating 8% of the population when the worms are highly clumped ($k=0.05$), while 52% of the population must be treated to achieve the same result when the parasites are randomly distributed ($k\to\infty$). (2) There is little benefit to be gained from being too selective (Fig. 5*b*, *d*, *f*). This is fortuitous since it would be difficult in practice to distinguish between individuals who carried say 50 or 75 worms. (3) Of all the three epidemiological variables considered in Fig. 5, selective treatment has the least effect on prevalence. The proportional reduction in prevalence declines as the selectivity index, *I*, increases (Fig. 5*d*).

The principle conclusion of this analysis is that, despite the extra cost involved in identifying heavily infected individuals, very substantial benefits may be gained from selective treatment, in terms of reducing both the abundance of the parasite (and hence, in general, the frequency of disease symptoms) and the number of treatments administered. This is primarily a consequence of the highly clumped nature of helminth distributions in human communities. The analysis is most applicable to the control of gut-dwelling helminths. For filarial infections, a reduction in worm burden (and hence transmission) may have no short-term impact on disease prevalence as a result of the permanent damage (such as blindness) induced by heavy infections before control. An added problem in the case of tissue-dwelling helminths is that drug-induced parasite mortality may have toxic side effects in those individuals harbouring heavy worm burdens.

Control by vaccination

Although vaccines against helminth infections are not available at present[11,53], we briefly consider the potential benefits to be gained, at the community level, from their development and use. If a safe and effective vaccine is developed, giving protection for an average period of *v* years, then the proportion of the population, *p*, that must be immunized per unit of time to reduce *R* below unity must exceed a critical value $\bar{p}$:

$$\bar{p}=[1-(1/R_0)]v^{-1} \qquad (17)$$

For example, if the vaccine gives life protection against a helminth infection with an R_0 value of 3 (appropriate for hookworm in endemic areas, see Table 2), it would be necessary to protect 67% of the community by a single vaccination (or course of vaccination) in infancy. However, if the vaccine only gives protection for a few years, then a large proportion of the population must be repeatedly immunized to sustain the necessary level of herd immunity for community protection. The relative merits of this type of vaccine when compared with cheap antihelmintic agents will therefore depend critically on cost factors since both forms of treatment will have to be administered repeatedly.

Conclusions

A major priority for the community control of infectious diseases is clearly the development of safe, effective and cheap agents (whether drugs or vaccines) to protect the individuals most at risk. These goals have been largely achieved for many of the common infectious diseases and yet they remain endemic throughout large regions of the world[49]. In the less developed countries, the availability of resources for primary health care is a major factor restricting the effective control of these diseases. It would therefore seem desirable to devote effort to elucidating the optimum ways of using the available methods,

© 1982 Nature Publishing Group

for the benefit of the community as a whole. A knowledge of the population dynamics of the disease agent has a central role in this field of research[28,30,33].

Human helminth infections are good examples of the discrepancy between our knowledge of how to treat an individual and how to control the infection in a community. The present analysis suggests that it is possible to determine, in quantitative terms, the level of antihelmintic treatment required in the community either to suppress parasite abundance to a defined level (to eliminate disease symptoms) or to eradicate the infection. More important, given the limitation of resources in areas where such infections are endemic, the analysis indicates that more attention should be given to the practical problems involved in selective treatment because the benefits to be gained by this approach seem very substantial. These benefits will be further enhanced, if, as has been suggested (but as yet unconfirmed[54]), the heavily infected individuals are predisposed to this state, not by the laws of chance, but as a consequence of genetic, behavioural or social factors.

This work was supported in part by US NSF grant DEB81-02783 (R.M.M.) and in part by the UNDP/World Bank WHO Special Programme for Research and Training in Tropical Diseases (R.M.A.). We thank K. Dietz for helpful discussions and comments on the manuscript.

Received 5 January; accepted 18 April 1982.

1. Stoll, N. R. *J. Parasit.* **33,** 1–18 (1947).
2. Iarotski, L. S. & Davis, A. *Bull. Wld Hlth Org.* **58,** 115–127 (1981).
3. Walsh, J. A. & Warren, K. S. *New Engl. Med. J.* **301,** 967–974 (1979).
4. Peters, W. in *The Relevance of Parasitology to Human Welfare Today* (eds Taylor, A. E. & Muller, R.) (Blackwell Scientific, Oxford, 1978).
5. Warren, K. S. *A. Rev. pub. Hlth* **2,** 101–115 (1981).
6. Anderson, R. M. in *Population Dynamics* (eds Anderson, R. M., Turner, B. D. & Taylor, L. R.) 245–282 (Blackwell Scientific, Oxford, 1979).
7. Anderson, R. M. *Lect. Notes Biomath.* **39,** 278–322 (1980).
8. Muller, R. *Worm and Disease: A Manual of Medical Helminthology* (Heinemann, London, 1975).
9. Warren, K. S. & Mahmoud, A. A. F. (eds) *Geographic Medicine for the Practitioner* (University of Chicago Press, 1978).
10. Schad, G. A. & Rozeboom, L. E. *A. Rev. Ecol. Syst.* **7,** 393–420 (1976).
11. Smith, M. A., Clegg, J. A., Snary, D. & Trejdosiewicz, A. J. *Parasitology* **84,** 83–92 (1982).
12. Van den Bossche, H. *Nature* **274,** 626–630 (1978).
13. W. H. O. *Tech. Rep. ser.* No. 615 (1977).
14. Anderson, R. M. & May, R. M. *Nature* **280,** 361–367 (1979).
15. May, R. M. & Anderson, R. M. *Nature* **280,** 455–461 (1979).
16. Croll, N. A., Anderson, R. M., Gyorkos, T. W. & Ghadirian, E. *Trans. R. Soc. trop. Med. Hyg.* **76,** 187–197 (1982).
17. May, R. M. *Math. Biosci.* **35,** 301–343 (1977).
18. Nawalinski, T., Schad, G. A. & Choudhury, A. B. *Am. J. trop. Med. Hyg.* **27,** 1152–1161, 1162–1173 (1978).
19. Wakelin, D. J. *Nature* **273,** 617–620 (1978).
20. Mitchell, G. F. *Adv. Immun.* **28,** 451–511 (1979).
21. Anderson, R. M. & May, R. M. *J. Anim. Ecol.* **47,** 219–248 (1978).
22. May, R. M. & Anderson, R. M. *J. Anim. Ecol.* **47,** 249–267 (1978).
23. Anderson, R. M. *J. theor. Biol.* **82,** 283–311 (1980).
24. Crofton, H. D. *Parasitology* **62,** 179–193 (1971).
25. Macdonald, G. *Trans. R. Soc. trop. Med. Hyg.* **59,** 489–506 (1965).
26. Bradley, D. J. & May, R. M. *Trans. R. Soc. trop. Med. Hyg.* **72,** 262–273 (1978).
27. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
28. Dietz, K. in *Population Dynamics of Infectious Diseases: Theory and Applications* (ed. Anderson, R. M.) (Chapman & Hall, London, in the press).
29. Dietz, K. *SIAM-SIMS Conf. Ser.* **21,** 104–121 (1975).
30. Anderson, R. M. & May, R. M. *Science* **215,** 1053–1060 (1982).
31. Anderson, R. M. (ed.) *Population Dynamics of Infectious Diseases: Theory and Applications* (Chapman & Hall, London, in the press).
32. Anderson, R. M. in *Modern Parasitology* (ed. Cox, F. E. G.) (Blackwell Scientific, Oxford, in the press).
33. Dietz, K. in *Population Biology of Infectious Diseases* (eds Anderson, R. M. & May, R. M.) (Springer, Berlin, in the press).
34. Bradley, D. J. & McCullough, F. S. *Trans. R. Soc. trop. Med. Hyg.* **67,** 491–500 (1973).
35. Muench, H. *Catalytic Models in Epidemiology* (Harvard University Press, 1959).
36. Warren, K. S. *J. infect. Dis.* **127,** 595–609 (1973).
37. Smithers, S. R. (ed.). *New Developments with Human and Veterinary Vaccines* 287–299 (Liss, New York, 1980).
38. Wagner, W. H. *Interdisc. Sci. Rev.* **5,** 186–203 (1980).
39. Chavania, A. P., Swalzelder, J. C., Villarejos, V. M. & Zeledon, R. *Am. J. trop. Med. Hyg.* **22,** 592–595 (1973).
40. Willett, W. C., Kilama, W. L. & Kihamia, C. M. *Am. J. Pub. Hlth* **69,** 987–991 (1979).
41. Janckes, M. F., Cornet, P. & Thunpont, D. *Trop. geogr. Med.* **31,** 111–122 (1979).
42. Santos, A. *et al. Bull. Wld Hlth Org.* **57,** 793–799 (1979).
43. Davis, A. & Wiegner, D. H. G. *Bull. Wld Hlth Org.* **57,** 767–771 (1979).
44. Ishizaki, T., Kamo, E. & Boehme, K. *Bull. Wld Hlth Org.* **57,** 787–791 (1979).
45. Davis, A., Biles, J. E. & Ulrich, A. M. *Bull. Wld Hlth Org.* **57,** 773–779 (1979).
46. Warren, K. S. in *Principles and Practice of Infectious Diseases* (eds Mandell, G. L., Douglas, R. G. & Bennett, J. E.) (Wiley, New York, 1979).
47. Arfaa, F. & Ghadirian, E. *Am. J. trop. Med. Hyg.* **26,** 866–871 (1977).
48. Smillie, W. G. *S. med. J., Nashville* **17,** 414–499 (1924).
49. Rougement, A., Boisson, M. E., Borges de Silva, G. & Zander, N. *Bull. Wld Hlth Org.* **54,** 403–410 (1976).
50. Warren, K. S. & Mahmoud, A. A. F. *Trans. Ass. Am. Physns* **89,** 195–204 (1976).
51. Warren, K. S., Siongok, A., Houser, T. K., Ouma, J. H. & Peters, P. A. *J. infect. Dis.* **138,** 849–855; **138,** 856–858 (1978).
52. Araujo, N., Katz, N., Pinto Dias, E. & Pereira de Souza, C. *Am. J. trop. Med.* **29,** 890–894 (1980).
53. Lloyd, S. *Parasitology* **83,** 225–242 (1981).
54. Croll, N. A. & Ghadirian, E. *Trop. geogr. Med.* **33,** 241–284. (1981).
55. Ansari, N. (ed.) *Epidemiology and Control of Schistosomiasis* (Karger, Basel, 1973).
56. Anderson, R. M. & May, R. M. *Parasitology* **79,** 63–94 (1979).
57. Anderson, R. M. in *Vectors of Disease Agents, Interactions with Plants, Animals and Man* (eds McKelvey, J. J., Elridge, B. F. & Maramorosch, K.) 13–43 (Praager, New York, 1981).
58. Krafsur, E. S. & Garrett-Jones, C. *Trans. R. Soc. trop. Med. Hyg.* **71,** 155–160 (1977).
59. Cheever, A. W. *Am. J. trop. Med. Hyg.* **17,** 38–64 (1968).
60. Hesieh, H. C. *Jap. J. Parasit.* **19,** 508–522 (1970).
61. Pesigan, T. P. *et al. Bull. Wld Hlth Org.* **18,** 345–455 (1958).
62. Bhallacharya, N. C. & Gubler, D. J. *Indian J. med. Res.* **61,** 8–11 (1973).
63. Hill, R. B. *Am. J. Hyg.* **6,** 19–41 (1926).
64. Stoll, N. R. *Am. J. Hyg.* **3,** 103–117 (1923).
65. Cheever, A. W., Kaner, I. A., Elwi, A. M., Moismann, J. E. & Danner, R. *Am. J. trop. Med. Hyg.* **26,** 702–716 (1977).
66. Miller, M. J., Krupp, I. M., Little, M. D. & Santos, C. *J. Am. med. Ass.* **230,** 1412–1414 (1974).
67. Wolfe, M. S. & Wershing, J. M. *J. Am. med. Ass.* **230,** 1408–1411 (1974).
68. Cameron, T. W. M. *Parasites and Parasitism* (Methuen, London, 1956).
69. Mahmoud, A. A. F. & Warren, K. S. *Clin. Res.* **28,** 474 (1980).
70. Dietz, K. *Lect. Notes Biomath.* **11,** 1–15 (1976).

Structure of the expanded state of tomato bushy stunt virus

I. K. Robinson & S. C. Harrison

Gibbs Laboratory, Harvard University, Cambridge, Massachusetts 02138, USA

The full three-dimensional structure determination of tomato bushy stunt virus has shown how its protein subunit can accommodate to different packing environments in the viral coat and how the size of the particle is nonetheless determined without ambiguity. Reversible, cooperative expansion of the virus at pH *values above neutrality and in the absence of divalent cation requires further bonding alternatives for the subunit. We report here the structure of the expanded tomato bushy stunt virus particle, determined at 8 Å resolution by X-ray crystallography. Subunit conformational changes are localized, and folded domains are essentially invariant. The bonding properties of the subunit can account for controlled polymorphism of the assembly.*

THE tomato bushy stunt virus (TBSV) particle comprises a single strand of RNA (molecular weight (M_r) 1.5×10^6) and 180 copies of a 40,000 M_r coat protein[1]. The subunits of the coat are arranged in a $T=3$ icosahedral lattice[2] as shown in Fig. 1*a*. High-resolution crystal structure analysis[3] has revealed the positions of all amino-acid residues in the C-terminal 75% of the protein chains; these form a shell having inner and outer radii of 110 and 170 Å. The N-terminal parts of the protein subunits and the whole of the RNA are invisible in the icosahedrally averaged, high-resolution electron density maps; they

0028–0836/82/240563—06$01.00

© 1982 Macmillan Journals Ltd

are thus believed to be spatially disordered. These components together fill the cavity in the interior of the virus[4,5].

The ordered part of the protein folds into two structural domains linked by a short 'hinge'. The 170-amino acid 'S'-domains pack tightly into the shell between radii of 110 and 140 Å. The three symmetrically distinct positions of this domain, shown in Fig. 1*b*, are denoted A, B and C. They have identical folding of their polypeptide backbone, but adjust to three 'quasi-equivalent' environments by differences in conformation of certain side chains on their surfaces and, more dramatically, by the ordered folding of 35 residues at the N-terminal end of the 60 C-position S-domains. These 35 residues form the so-called 'extended arm' and 'β-annulus'[3]. The A- and B-position S-domains make an extensive contact with each other, which is broken in C positions by interposition of the arm. Distal residues of the arm interdigitate with corresponding residues of two other C subunits to form the β-annuli. A coherent internal scaffold is thereby created, which precisely determines the curvature and size of the virus particle.

The second domain, called 'P', forms strong dimers that lie on the 2-fold and quasi-2-fold axes of the constellation of S-domains. These dimers project to the outer radius of 170 Å and are clearly visible in electron micrographs of the virus[6].

Expansion

TBSV and many other related viruses undergo a structural phase transition in solution that is controlled primarily by *p*H and divalent cation concentration. The phenomenon was first observed in bromegrass mosaic virus (BMV) by Incardona and Kaesberg[7] using various physicochemical techniques. These all gave results consistent with a model in which the particle mass was conserved and the hydrodynamic radius increased by ~10%. For this reason we will refer to it as 'expansion' of the virus.

BMV exhibits an anomalous titration of two protons per subunit associated with the phase transition at *p*H 6.7. This has been interpreted[8,9] as due to formation of hydrogen bonds

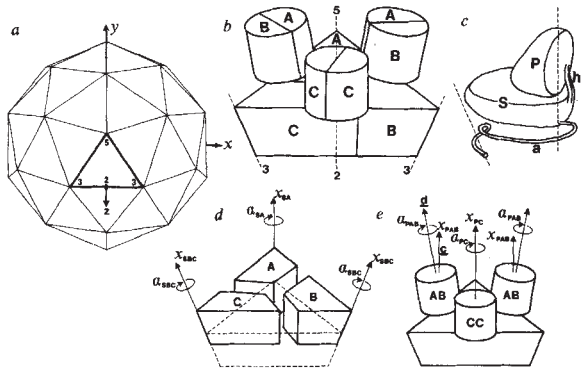


Fig. 1 Architecture of the TBSV particle as determined by high-resolution X-ray structure analysis, and model of its expansion. *a*, Stellated pentagonal dodecahedron showing the definition of the icosahedral asymmetric unit (highlighted) as the spherical sector lying between a 5-fold and two adjacent 3-fold axes. The viral shell consists of 60 of these identical units. *b*, Schematic diagram showing the packing of three quasi-equivalent TBSV protein subunits in the icosahedral asymmetric unit, denoted A, B and C. *c*, Approximate shape of the C-position subunit, showing the two domains, denoted 'S' (surface) and 'P' (projecting), connected by a polypeptide hinge, h. The N-terminal arm, a, is ordered only in the C-position subunit and disordered in A and B. *d*, Parameterization of the model of the expansion of the S-domains. A translation, *x*, parallel to the symmetry axis and a rotation, α, about that axis is the most general description. The B- and C-position domains are assumed to retain their contacts about the 3-fold axis and so are described by the same parameter pair. *e*, Parameterization of the P-domain expansion. The AB P-domain dimer is unconstrained by the particle symmetry, and so has the full six parameters of a rigid body transformation. Spherical polar coordinates are used for compatibility with the other parameter definitions.

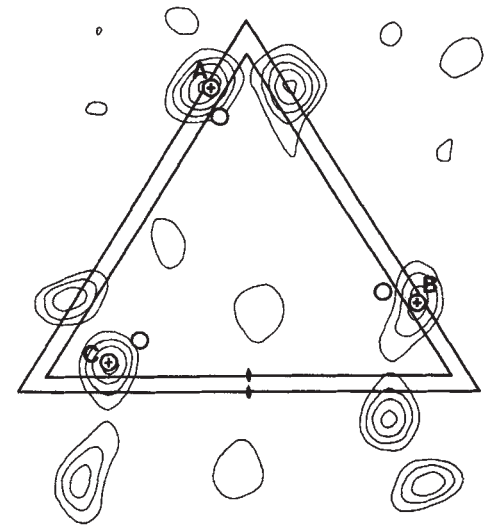


Fig. 2 A section of the 12 Å difference Fourier for the $PtCl_4$ derivative, located 132 Å from the particle centre. Phases for this map were derived from the best three-parameter model and refined by symmetry averaging. The larger triangle denotes the boundary of the icosahedral asymmetric unit and the symbol ⊕ denotes the refined atomic positions. The quasi-symmetry of these site positions is clearly apparent. Superimposed on the same scale are the refined platinum positions for the compact virus (○), which lie at a radius of ~116 Å, and the size of the icosahedral asymmetric unit at that radius (smaller triangle). The vector joining the platinum position in the compact virus to that of the expanded virus is an accurate measure of the displacement of the subunit to which the atom is bound.

between apposed carboxylic acid groups. Divalent cations, known to inhibit the expansion of BMV, would displace the hydrogen bonds in favour of ionic links stable at all *p*H values. Native preparations of TBSV have two bound, EDTA-chelatable ions per subunit (probably both calcium) which must be removed before expansion can be observed[10]. Once free of these ions, the virus expands reversibly when the *p*H is raised above 7.0; the transition seems to be characterized by some hysteresis, which is most easily followed by changes in the solution X-ray scattering. The location and structure of the cation binding sites[11] are consistent with the mechanism just described: five aspartate side chains, contributed by each pair of adjacent S-domains in the trimer of Fig. 1*b*, are involved in the binding of the two ions. When charged at high *p*H, these aspartates repel each other and destabilize the inter-domain contact. This interface is therefore likely to be the one broken to permit expansion.

Examination of the relative extent of various contacts between S-domains in the high-resolution structure[3,11] led to a model for expansion, shown in Fig. 1*d*. This model assumed that the virus retained its icosahedral symmetry and that the internal tertiary structure of each domain was conserved. Rapid reversibility of expansion further suggested that the interlocked internal scaffold was conserved and hence that the pentamer and hexmer clusterings of S-domains in the virus were unaltered. As argued above, it was reasonable to suppose that the trimer contacts, which contain the divalent cation sites, were disrupted. With this model, S-domain displacements can be described by just four parameters, two rotations about and two translations along the symmetry axes. The motions of the P-domains described in Fig. 1*e* are rather more general: only the PP dimer contact is conserved, so eight parameters make a complete description. We have used this model-building approach in the crystal structure determination.

Structure determination

Crystals of the expanded virus were prepared by vapour diffusion[12] from 2.9% polyethylene glycol, 0.01 M EDTA, 0.058 M $(NH_4)_2HPO_4$, 20% ethylene glycol, *p*H 7.5 at 18 °C.

© 1982 Nature Publishing Group

X-ray diffraction showed crystalline order to ~7 Å resolution. The space group is C2, with cell constants $a = 546.3$ Å, $b = 433.1$ Å, $c = 383.4$ Å and $\beta = 134.0°$. The crystallographic asymmetric unit contains one half of a virus particle, so there is 30-fold non-crystallographic symmetry, which can be exploited for refinement of phases[13]. The particle position within the unit cell is defined in this space group with the exception of the particle orientation angle about the 2-fold axis, referred to as θ, which must be determined experimentally; an additional parameter must therefore be added to the model description of the expanded virus in the crystals.

Data were collected by oscillation photography[14] for native crystals as well as $PtCl_4^{2-}$ and ethyl mercury thiosalicylate (EMTS) heavy-atom derivatives. These heavy-atom reagents were similar to those used in the structure determination of the compact virus[15], in the hope that the same sites would be labelled and a direct comparison of coordinates could be made. The data quality deteriorated sharply with resolution, and only reflections in the range 8–35 Å were considered usable. Table 1*a* summarizes native and derivative data collection. Comparison of Wilson plots for data from the compact and expanded forms of the virus indicated that the absence of high-resolution data for the latter was well modelled by gaussian disorder, with a Debye 'temperature factor' of 340 Å^2.

Starting values of the model parameters were obtained by analysing the packing of the expanded virus particles into the crystal unit cell. These were used to compute a predicted diffraction pattern, which was compared with the observed pattern by calculating an *R*-factor for agreement of scaled intensities. The model parameters were adjusted until the *R*-factor was minimized. A value of $R = 0.52$ was obtained for 12 Å, significantly lower than the value of 0.59 obtained by comparing sets of uncorrelated intensities. A more detailed account of this critical step and of the rest of the structure determination will be published elsewhere.

Phases calculated from the best model were applied to the observed structure facors to 12 Å resolution and subjected to refinement by non-crystallographic symmetry averaging[13]. The *R*-factor for comparison of symmetry-averaged and observed intensities was reduced to 0.38 in six cycles (12 Å data) and the phases on average moved by 55°.

To confirm the validity of the phasing and to extend its resolution, heavy-atom difference maps were calculated for both derivatives using the refined 12 Å phases. An example is shown in Fig. 2. Sites corresponding to the three platinum positions[15] were unambiguously identified. Their parameters were refined by conventional methods[16]. The *a priori* assumptions of rigid-body domain motion and conservation of hexamer and pentamer subunit contacts were supported by the finding that the positions of the S-domain heavy-atom sites were self-consistent to within 2 Å (Table 1*b*). The refined positions allowed a considerably more precise set of model parameters to be derived directly.

The three steps of model-building, *R*-factor minimization and non-crystallographic symmetry phase refinement were then repeated to extend the phasing to the full 8 Å data set. The refined heavy-atom occupancies at 8 Å were unchanged from their values at 12 Å (Table 1*b*), indicating the absence of significant deviations from icosahedral symmetry. Fourier analysis of the best model gave an *R*-factor of 0.46 with all the observed data. Phase refinement brought this value down to 0.43; the lack of dramatic improvement here indicates that the model description is a good one and reflects the poorer quality of the higher-resolution data. Electron density maps were prepared with dots superimposed for every amino acid residue of the model structure. Corresponding 8 Å maps of the compact structure with an imposed 340 Å^2 'temperature factor' were also prepared for comparison.

Interpretation of electron density

The successful refinement and correct determination of heavy-atom positions confirm our assumption that tertiary structure in each domain is conserved, with conformational changes restricted to localized flexible regions. The displacements and rotations generated by these local flexions are described by the refined model parameters in Table 2. A striking aspect of the expanded state is the appearance of a branched opening 80 Å long and large enough for a 20 Å sphere to pass through each of the 60 faces; this is seen in Fig. 3, and an overall view of the icosahedral asymmetric unit is shown in Fig. 4*a*. The P-domains are rotated by relatively large angles, 30° in the AB case and 103° for the CC dimer. This latter value is so great that the actual direction of the rotation is ambiguous; the clockwise direction is chosen on the basis of a more reasonable conformation of the hinge. To accommodate these rotations and to bridge the gap between the domains, it is necessary to postulate a new 'expansion hinge', several residues long, to the N-terminal side of the hinge associated with the quasi-symmetry[3]. We have thus seen four distinct states of the linkage

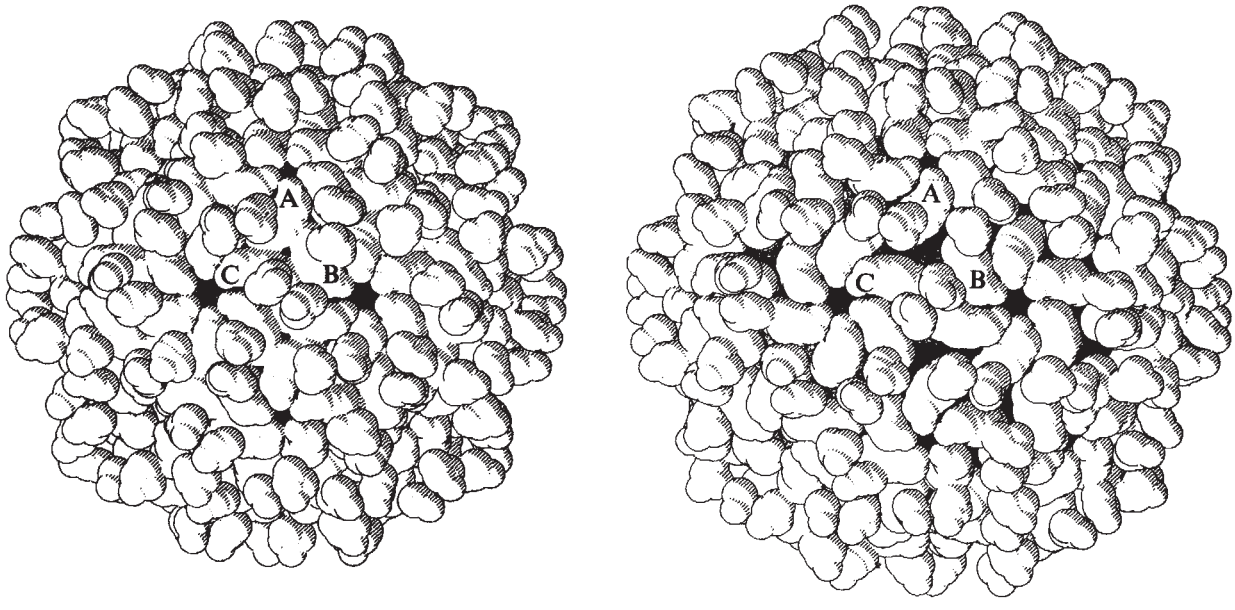


Fig. 3 Pictorial view of the expansion of TBSV. The compact particle is on the left. Note the appearance of the inter-subunit opening and the rotation of the projecting domains in the expanded virus on the right. (This figure was prepared by A. J. Olson using a modified form of the PLUTO program of S. Motherwell.)

© 1982 Nature Publishing Group

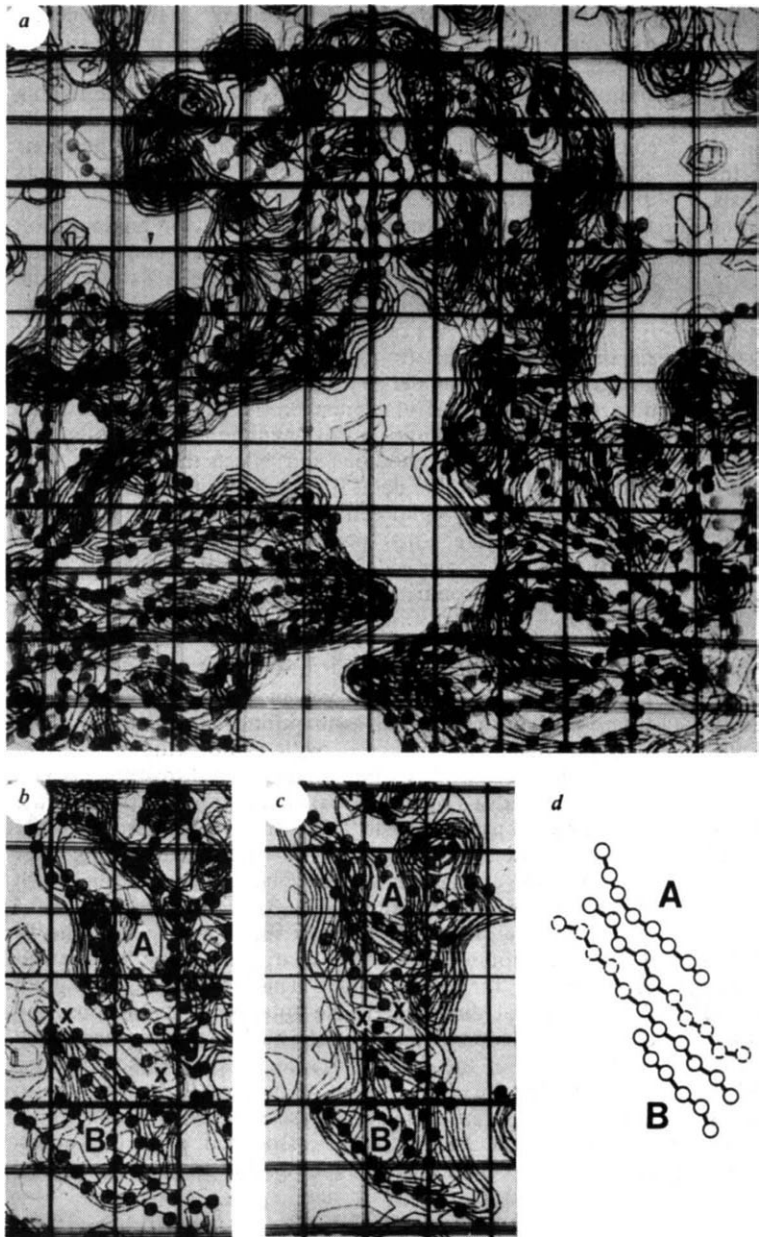


Fig. 4 Electron density maps at 8 Å resolution and their interpretation. Dots are superimposed for each amino acid in the best-fitting model of the expanded virus. The grid is 10 Å square. *a*, Overall view of the icosahedral asymmetric unit of expanded TBSV in the S-domain region. The opening is clearly visible between the domains. An impression of the quality of the map can be obtained from the fit of model to density and from the noise level in the solvent regions. *b*, The contact between the A and B position S-domains of the compact structure. The residue marked 'X' in each subunit is the first residue to be included at the N-terminus of the model. This part of the structure contains a four-strand antiparallel β-sheet extending across the base of the domain. *c*, The corresponding region of the expanded structure in which the two domains are shifted past each other by 14 Å. There is clear density immediately preceding residue X which can be interpreted as an ordering of the arm. *d*, Interpretation of the new electron density as an additional strand of β-sheet that forms a clamp between the subunits.

region between S- and P-domains: in A/B and in C positions in each of the compact and expanded particles.

The refined expansion parameters and the known high-resolution structures of the individual domains define an accurate model of the expanded virus. Use of this model greatly facilitated further interpretation of the refined electron density. Even though the resolution was not sufficient to distinguish individual strands of protein, the fit of model to density was good enough for small changes from the model to be detected reliably.

The extended arm, which is ordered in C-position S-domains but not in A or B, is conserved in the expanded virus, as is the β-annulus. The latter feature is much stronger in the expanded form; it appears ordered for several residues further into the interior of the virus. The function of the C-position arm and the β-annulus is to determine the size of the virus during assembly[3,11]. Conservation of these features in the expanded state is consistent with their fundamental role in establishing structural coherence.

The contact between the four-stranded β-sheets at the base of the A- and B-position S-domains is sheared apart by the expansion, as shown in Fig. 4*b* and *c*. A very distinct bulge of new density lies beside the first strand in both the A and B positions of the expanded state. This seems to be additional ordered protein at the N-terminus of the domain, and we suggest the interpretation of Fig. 4*d*. About six residues immediately N-terminal to the A and B position, S-domains of the compact structure become ordered in the expanded state, extending the first strand of each β-sheet and forming hydrogen bonds with residues of the opposite subunit. This interaction would create a single large sheet extending across the inner surface of both S-domains (Fig. 5). The density is interpreted in this way as the new B-position arm extending beyond the base of the A-position S-domain for a further 20 Å along the inside edge of the inter-subunit opening, terminating at a radius of 156 Å, close to the exterior of the virus. Thus, one of the arms seems to pass all the way through the opening to the outside, possibly explaining the observation that mild proteolysis of expanded virus causes a cleavage of many of the chains at positions 70 to 100 residues from the N-terminus[17].

A feature of the S-domain also seems to have shifted during expansion. The trimer contact region in the centre of the triangular icosahedral asymmetric unit of the virus (see Fig. 1*b*) consists of three loops of extended chain in symmetrical contact. In the maps of expanded virus, the rigid model of the domain has no electron density in this region, while a new loop of density of equal contour length protrudes into the opening. We suggest that a local refolding of the chain occurs in this region, with a displacement of ~10 Å at the tip of the loop. A similar

© 1982 Nature Publishing Group

change in secondary structure of an exterior loop has been observed in trypsinogen[18].

Apart from this limited number of modifications, which are easily rationalized, the model of the expanded virus is completely consistent with the electron density. As all the density is accounted for, we are confident of the interpretation.

Stability and disorder in the expanded structure

The detailed interpretation of electron density described above shows that the expanded state is characterized by a relatively rigid network of S-domains. This network may be stabilized in part by a new contact between the bases of the A and B S-domains, involving an ordered conformation for about six residues of the arm in those positions where it is not ordered in the compact state. The extended arm and β-annulus of the C-position subunits are conserved. The P-domains are rearranged on the surface of the virus by fairly large rotations about their local diad axes.

Table 1 Data collection statistics for expanded TBSV (*a*) and refined heavy-atom coordinates (*b*)

a

Resolution limit (Å)	16	12	8	
No. of possible reflections	6,888	17,323	60,193	
No. observed	3,880	10,356	36,113	Native
R_{sym}	0.10	0.12	0.20	
No. observed	508	1,357	4,733	$PtCl_4$
R_{sym}	0.10	0.12	0.16	
No. observed	1,351	3,607	12,579	EMTS
R_{sym}	—	0.11	0.20	

b

	x	*y*	*z*	Occupancy (12 Å)	Occupancy (8 Å)
Platinum					
A	−8	67	134	8.1	8.0
B	37	19	133	6.4	7.2
C	−30	7	133	6.7	7.0
Mercury					
A1	−17	44	146	4.1	4.2
B1	19	40	145	4.5	2.9
Cl	−5	9	145	3.2	2.7
A2	−29	48	178	2.2	4.5
B2	29	46	179	3.8	4.6
C2	10	5	179	2.1	2.2
E	−43	1	151	3.2	3.1

a, A total of 56 1° oscillation films were used in the final native data set. The reproducibility of the data is indicated for the redundant measurements by R_{sym}.

$$R_{sym} = \frac{\sum_h \sum_{j=1}^{n_h} |I_{hj} - \bar{I}_h|}{\sum_h n_h \bar{I}_h}$$

where the I_{hj}, $j = 1, \ldots n_h$ are the repeated measurements of I_h, and $\bar{I}_h$ is their mean. *b*, Refined heavy-atom coordinates (Å in coordinate frame of Fig. 1*a*) and relative occupancies (arbitrary units). Individual heavy-atom temperature factors were set to 25 Å^2 and not refined. Occupancies are shown for refinements at 12 Å and 8 Å resolution. All sites except E have chemically equivalent positions in the derivatives of compact TBSV, and their nomenclature corresponds to that of Table 2 in ref. 15. Site E is unique to the expanded structure and lies near a known cysteine.

Table 2 Final parameters for the best model of expanded TBSV

Translations		Rotations	
x_{SA}	19.4 Å	α_{SA}	4.5°
x_{SBC}	17.8 Å	α_{SBC}	2.1°
x_{PAB}	18.1 Å	α_{PAB}	30°
x_{PC}	14.2 Å	α_{PC}	103°

The parameters are defined in Fig. 1*d* and *e* legends. Direction cosines for AB P-domain dimer translation and rotation. $\vec{c}$, (−0.187, 0.352, 0.917); $\vec{d}$, (−0.264, 0.140, 0.954); θ, 97.1°.

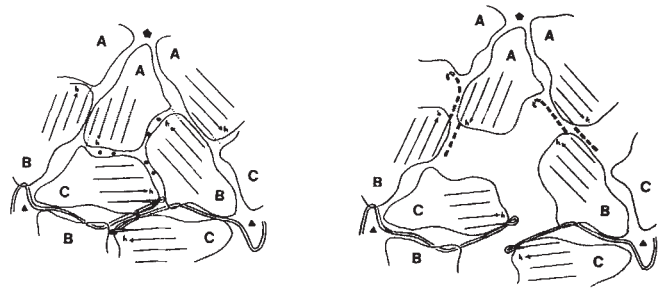


Fig. 5 Schematic diagrams showing the inner surface of S-domains in compact and expanded TBSV and indicating the interpreted disposition of the A and B position arms in the expanded structure. The Ca^{2+} sites are shown by pairs of dots[10]. The β-sheets at the base of the S-domains are denoted by four parallel lines, the C arms by double lines and the interpretation of A and B arms in expanded TBSV by broken lines. In this interpretation, the A arm extends to make a six-residue contact with the B subunit, and the B arm extends for about 15 residues, making the same 6-residue contact with the A subunit and continuing up the inside of the opening to the outside of the virus. The C arm has exactly the same conformation in both structures, except that in expanded virus it is ordered ~10 Å further into the viral interior. The strands terminating in arrows and indicated by 'h' lead to the inter-domain hinge. The change in spacing of these arrowheads across dyads when the particle expands illustrates the requirement for an 'expansion hinge' as described in the text.

The CC P-domain dimer appears to be more disordered than the rest of the structure: the corresponding heavy-atom site refined rather poorly, and the resulting electron density of this domain is considerably weaker than that of the quasi-symmetry-related AB dimer. A model was built with individual relative isotropic temperature factors for the various domains and optimized with respect to the observed dta. The best model had $B = 80$ Å^2 for the A and B position P-domains and $B = 600$ Å^2 for the C position. The latter temperature factor corresponds to a r.m.s. positional fluctuation relative to the S-domains of 5 Å (ref. 19), and because this domain is involved in the inter-particle contact of the crystals it can explain the disorder of the crystals as a whole.

Because of the disorder of the P-domains, it seems unlikely that they are involved in a rigid way with clamping the particle in the expanded state. Moreover, southern bean mosaic virus (SBMV), which has a single domain structurally homologous to the TBSV S-domain and no P-domain[20], also has an expanded particle[21]. If the structural similarity of these viruses extends to the expanded state, the P-domain cannot be specifically involved in the expansion mechanism, and its motions must be a consequence of more fundamental structural changes occurring elsewhere in the virus. The proposed AB S-domain contact (Fig. 5) is a feature that could equally well exist in the expanded state of SBMV as in TBSV.

There is no evidence for ordering of the RNA or the N-terminal portion of the protein, other than the places mentioned. The inter-subunit opening is sufficiently large for the arm to be able to extrude entirely from the virus, even if the N-terminal 65 residues were folded as a third domain of the structure. That some arms do extrude is consistent with their proteolytic sensitivity in expanded TBSV[17].

The expansion mechanism is triggered by deprotonation of the aspartate residues of the calcium binding sites, provided the Ca^{2+} salt bridges have been removed by chelation (Fig. 5). The local build-up of charges must then prise the opening apart. The repulsive force, which would increase monotonically with *p*H in the absence of any induced structural change, is counterbalanced by the attractive forces between domains and between the RNA and the basic amino acids of the interior surface. Beyond a critical *p*H, the balance swings in favour of expansion, and the particle switches to its other state by rearrangement of

© 1982 Nature Publishing Group

the S-domains. The P-domains then presumably follow suit by adopting minimum-energy configurations on the new surface.

Functions of the structural change

The physiological significance of expansion, if any, is unknown. The expanded state, unlike the compact one, is very sensitive to enzymatic proteolysis, just as it is for BMV[22]. In both cases, this seems to be due to accessibility of normally buried N-terminal sequences. These common features may reflect a common route of entry and disassembly, and Durham[23] has suggested that a difference in calcium concentration between the cytoplasm and the extracellular fluid might be significant in this regard. It is also possible that the expanded state is an intermediate in the virus self-assembly. Whatever its role, the expanded particle demonstrates three additional conformations of the viral subunit. As in the compact structure, the polymorphism is restricted to specific regions of the polypeptide that form specific links between globular domains.

The expansion of TBSV is a highly cooperative change from one state of the viral shell to another: indeed, it has the character of a first-order phase transition[17,24]. There is restructuring of the 2-fold contacts, with conservation of the 5- and 6-fold contacts, giving a two-state character to the transformation. Observed deviations from formal two-state behaviour[17,24] are probably due to changes in the spatially disordered part of the structure, particularly in the N-terminal area. We have strong evidence from the cited proteolytic cleavage experiments[17] that some of these arms extend outside the expanded particle in certain conditions. The manner in which restructured non-covalent contacts are built into the subunit design suggests that the expanded form has a function in the virus life cycle. Indeed, expansion seems to be a general property of small RNA plant viruses. There might also be some formal analogy with the reversiblly interconvertible isoelectric forms of small RNA animal viruses[25].

We thank J. M. Hogle, A. J. Olson, C. S. Steele and D. C. Wiley for advice, discussion and assistance. This research was supported by NIH grant CA-13202.

Received 14 October 1981; accepted 30 March 1982.

1. Ziegler, A., Harrison, S. C. and Leberman, R. *Virology* **59**, 509 (1972).
2. Caspar, D. L. D. & Klug, A. Cold Spring Harb. Symp. quant. Biol. **27**, 1 (1962).
3. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368 (1978).
4. Chauvin, C., Witz, J. & Jacrot, B. *J. molec. Biol.* **124**, 641 (1978).
5. Munowitz, M. G., Dobson, C. M., Griffin, R. G. & Harrison, S. C. *J. molec. Biol.* **141**, 327 (1980).
6. Finch, J. T., Klug, A. & Leberman, R. *J. molec. Biol.* **50**, 215 (1970).
7. Incardona, N. L. & Kaesberg, P. *Biophys. J.* **4**, 11 (1964).
8. Bancroft, J. B. *Adv. Virus Res.* **16**, 99 (1970).
9. Caspar, D. L. D. *Adv. Protein Chem.* **18**, 37 (1963).
10. Hogle, J. M., Kirchhausen, T. & Harrison, S. C. *J. molec. Biol.* (submitted).
11. Harrison, S. C. *Biophys. J.* **32**, 139 (1980).
12. McPherson, A. *Meth. Biochem. Analysis* **23**, 249 (1975).
13. Bricogne, G. *Acta crystallogr.* **A30**, 395 (1974).
14. Arndt, U. W. & Wonacott, A. J. *The Rotation Method in Crystallography* (North-Holland, Amsterdam, 1977).
15. Winkler, F. K., Schutt, C. E., Harrison, S. C. & Bricogne, G. *Nature* **265**, 509 (1977).
16. Dickerson, R. E., Kendrew, J. C. & Strandberg, B. E. *Acta crystallogr.* **14**, 1188 (1961).
17. Sorger, P. K., Stockley, P. G., Robinson, I. K. & Harrison, S. C. (in preparation).
18. Stroud, R. M., Kossiakoff, A. A. & Chambers, J. L. *A. Rev. Biophys. Bioengng* **6**, 177 (1977).
19. Debye, P. *Ann. Phys.* **43**, 49 (1914).
20. Abad-Zapatero, C. *et al. Nature* **286**, 33 (1980).
21. Hsu, C. H., White, J. A. & Sehgal, O. P. *Virology* **81**, 471 (1977).
22. Pfeiffer, P. & Hirth, L. *FEBS Lett.* **56**, 144 (1975).
23. Durham, A. C. H., Hendry, D. A. & von Wechmar, M. B. *Virology* **77**, 524 (1977).
24. Kruse, J. & Witz, J. (in preparation).
25. Rueckert, R. R. *Comp. Virol.* **6**, 131 (1976).

LETTERS TO NATURE

Discovery of 69 ms periodic X-ray pulsations in A0538−66

G. K. Skinner*, D. K. Bedford*, R. F. Elsner†, D. Leahy†, M. C. Weisskopf† & J. Grindlay‡

* Department of Space Research, University of Birmingham, Birmingham BI5 2TT, UK
† Space Sciences Laboratory, NASA Marshall Space Flight Center, Alabama 35812, USA
‡ Center for Astrophysics, 60 Garden St, Cambridge, Massachusetts 02138, USA

Both the recurrent X-ray transient A0538−66 and its optical counterpart undergo outbursts at intervals which are multiples of 16.6 days[1–3]. The times of outbursts are subject to appreciable jitter and although typically brief (<1 day) the event can, on occasions, last nearly a complete cycle[4]. The system lies in the direction of the Large Magellanic Cloud (LMC) and the Doppler shifts of the quiescent-state optical absorption lines indicate that it is a member of the LMC[5,6]. This implies an X-ray luminosity during outburst of 8×10^{38} erg s^{-1} (2–17 keV)[1]—higher than that of any other 'stellar' X-ray source. We report here the detection, during observations with the Einstein Observatory at the time of an outburst, of X-ray pulsations with a period of 69.2126 ms from A0538−66, adding to the unique properties of this system that of being the fastest X-ray pulsar known to be in a binary system. The observed rate of change of period shows the source to be an eccentric binary system in which the two components approach close to each other during the outbursts and may even become immersed in a common envelope.

Two Einstein observations of A0538−66 were made as part of the Guest Observer Program. One of these was made on 16 December 1980 23.73 to 24.85 h UT, at phase 0.056–0.058, where phase zero corresponds to the predicted time of a peak[7]. A point source coincident with the position of A0538−66 was detected by the High Resolution Imager (HRI) in the focal plane of the telescope at a mean level of (0.090±0.005) HRI counts s^{-1}. Unlike the HRI, the Monitor Proportional Counter (MPC) which is aligned with the main telescope and has a field of view of 42 arc min (FWHM) is sensitive to relatively high-energy X rays (1.1–21 keV) and provides energy resolution. The MPC registered a flux ~40 MPC counts s^{-1} above background. The spectrum of the source detected by the MPC showed a strong low energy cutoff corresponding to photoelectric absorption in a column density $\cong 2\times10^{23}$ cm^{-2} (using Brown and Gould[8] cross-sections) making this flux consistent with that observed from A0538−66 with the HRI. Simultaneous fluctuations on time scales 10–1,000 s in the two flux levels, apparently due to variations in the column density of absorbing matter, confirm that the MPC flux was from A0538−66.

The MPC includes a time interval processor (TIP) which records the time between events in the 1.1–21 keV energy band with a resolution of 1 μs or 1.6%, whichever is the greater. The TIP data were analysed by converting to photon arrival times, correcting for the motion of the satellite and of the Earth, and examining them for manifestations of variability. This procedure revealed the existence of a persistent periodicity, detected throughout the entire observation at a mean heliocentric pulse period of 0.0692126 s. Epoch folding the data at the best period and correcting for the existence of a period derivative leads to a result 111σ above the noise, while folding at neighbouring but statistically independent periods produces results consistent with noise.

The pulse period was determined by means of pulse arrival time analysis. Initially the entire data set was folded modulo different trial periods and the χ^2 statistic was maximized as a function of period. The resulting light curve was used as a template and each 208 s section of the data was compared with it using a cross-correlation technique, allowing a relative pulse arrival time, t_m, to be determined. The t_m values were fitted by functions of the form $t_m = t_0 + mP$ and $t_m = t_0 + mP + \frac{1}{2}m^2P\dot{P}$. The fit assuming a constant period was unsatisfactory ($\chi^2 = 52$

0028–0836/82/240568—03$01.00

© 1982 Macmillan Journals Ltd

for 16 d.f.). Figure 1*b* shows the residuals for this fit and also the curve corresponding to the fit obtained once a period derivative was included, which was excellent ($\chi^2 = 17.7$ for 15 d.f.). The parameters for the best fit obtained are: pulse period $P = (0.06921166 \pm 0.00000017)$ s, at $t_0 =$ JD 2444590.4792, rate of change of period $\dot{P} = (5.01 \pm 0.86) \times 10^{-10}$. All times are heliocentric. Errors are $\pm 1\sigma$.

Figure 1*a* shows the light curve obtained by folding the data on the basis of the best fit P, $\dot{P}$. After allowing for the background, the pulsed component accounts for 26% of the flux from the source. The amplitude of the pulsations in the separate 208s intervals follows the mean count rate above background so that there is no evidence for the pulsed-fraction changing; the pulsed flux is apparently affected by the variable absorption column density in the same way as the steady flux.

The second scheduled Einstein observation of A0538−66 took place on 3 February 1981 23.78–25.76 h UT at phase −0.002 to +0.004 (no data were available for 24.28–25.00 h UT). In both the HRI and MPC data A0538−66 was seen to be turning on during the last ~1,400 s, averaging 10 MPC counts s^{-1} during this time. No pulsations were detected although they would have been if the pulse fraction and shape had been the same as during the more intense first observation. However, the energy spectrum was considerably harder. The photon power law index was −0.6 during the second observation, compared with −1.6 during the first, although the average column densities were similar. We note that variations of pulse fraction and shape with intensity were found in Uhuru observations of Cen X-3 (ref. 9).

A0538−66 was within the field-of-view of the Einstein telescope on two other occasions: 10 April 1979 12.73–13.70 h UT (phase −0.002 to +0.002) and 30 May 1980 9.07–10.63 h UT (phase +0.024 to +0.026). Despite the fortuitous proximity to phase zero, in neither case was significant flux detected with either the MPC or the main telescope. This may be related to the apparent absence of any large scale activity during the optical observations of Pakull and Pamar[5] in 1980.

The relationship $P \gtrsim (G\bar{\rho})^{-1/2}$ between pulsation period and the mean density of the region in which it arises shows that the seat of the 69 ms pulsations must involve a neutron star (or a black hole). The observed $\dot{P}$ is smaller by a factor of $>10^3$ than

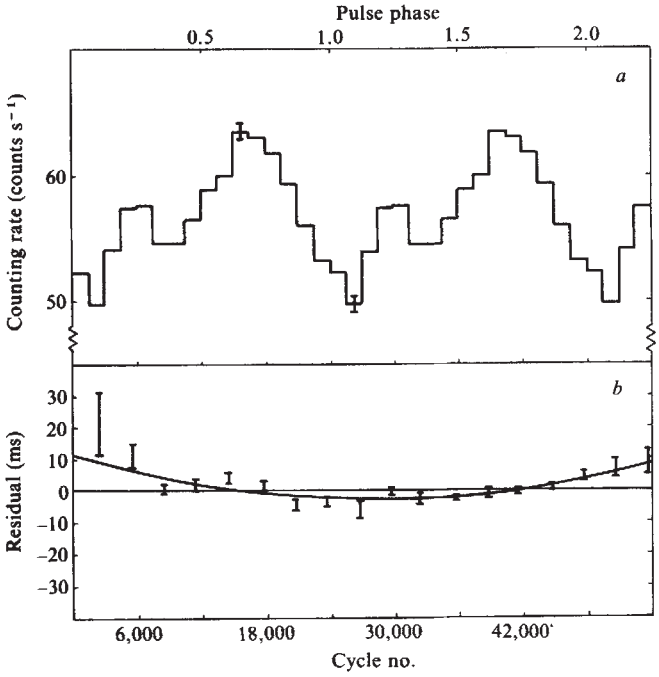


Fig. 1 *a*, The MPC data folded according to the best fit parameters. *b*, Measured pulse arrival time minus that expected on the assumption of a constant period $P = 0.06921264$ s. The curve corresponds to $P = (0.06921166 \pm 0.00000017$ s; $\dot{P} = 5.01 \pm 0.86 \times 10^{-10}$. Error bars are $\pm 1\sigma$.

that of the oscillations reported by Sadeh *et al.*[10] during a burst from MXB1728-34 and we believe that the system is a true pulsar in which the clock is a rotating neutron star.

If the observed $\dot{P}$ were to correspond to a changing rate of rotation, a power of $\sim 10^{41}$ erg s^{-1} would have to be dissipated in the system. As this seems unreasonable, and the expected time scale for spin-up or spin-down due to accretion torques is too long (see below), we conclude that the major component of the $\dot{P}$ term is due to a changing Doppler shift. Consequently, the neutron star must be in a gravitational field whose line of sight component causes an acceleration of 220 cm s^{-2} and, as optical evidence suggests that the primary has mass $M_{opt} = 12\,M_\odot$, the stellar separation can be no more than $39\,R_\odot$. Charles *et al.*[11] have shown that the region responsible for the optical flux during outburst must be at least $40\,R_\odot$ in radius and so, whether the optical emission is centred on the primary or the secondary, it is already clear that during the outburst we are dealing with two objects essentially in contact.

The maximum acceleration of the secondary in a binary system with period P_{orb} and eccentricity e is

$$(2\pi)^{4/3} G^{1/3} (1-e)^{-2} M_{opt}^{1/3} (1+q)^{-2/3} P_{orb}^{-4/3}$$

where q is the mass ratio M_X/M_{opt}. If we assume that the orbital period is 16.6 days, then we find $e \gtrsim 0.4$ (we have taken $M_{opt} = 12 M_\odot$, $q = 0.1$, although the conclusion depends only weakly on the choice). The limit $e = 0.4$ corresponds to the case where the observation was made exactly at periastron and the inclination and orientation of the orbit are such that the neutron star was directly between the observer and the primary. Any other configuration implies an even higher eccentricity.

While not losing sight of the possibility that the source of energy might be rotational, through a pulsar mechanism, we now consider the implications of powering the X-ray source by accretion onto the neutron star. The first question is whether accretion is possible in the presence of a magnetic field rotating with such a high angular velocity. Before accreting onto the neutron star, matter must pass through a co-rotating magnetosphere. If the radius of the magnetosphere is larger than that of the keplerian orbit co-rotating with the neutron star, accretion is unlikely. The size of the magnetosphere will depend not only on the magnetic field but also on the amount of inflowing matter and so on the X-ray luminosity. We deduce

$$B \lesssim f \times 10^{11} \left(\frac{L_X}{10^{39}\ \mathrm{erg\ s^{-1}}}\right)^{1/2} \left(\frac{P}{69\ \mathrm{ms}}\right)^{7/6} \quad \mathrm{G}$$

where f is a numerical factor ~0.5–1.0 depending on the detailed assumptions made and on whether the equations used are those for accretion from a wind[12,13], or for accretion through a disk[14]. We can envisage a situation where the magnetic field is close to this critical value and, as periastron is approached, essentially no accretion is possible until a sufficiently large amount of accreting matter is available to produce a high X-ray luminosity. This would explain the very large on/off ratios for the X-ray flux, the sharp turn-ons and turn-offs, and the small duty cycle of the on state. In this picture although the neutron star would spin-down between outbursts it might either spin-up or spin-down during the outburst. The contribution to the short-term $\dot{P}$ would in either case be small compared with that observed. We note that if the magnetic field were too small non-uniform accretion onto the newton star surface, necessary to explain the 69-ms modulation of the X-ray flux, would not occur.

We may consider whether a mechanism analogous to Roche Lobe overflow can be responsible for the accretion. To estimate the size of the companion's critical lobe at periastron, we adapted the results quoted by Bahcall[15] to the pseudo-potential given by Avni[16] describing the gravitational/centrifugal forces on a test particle co-rotating with the companion and in an eccentric binary system. Using the parameters found by Charles *et al.* for the quiescent state ($M_{opt} = 12 M_\odot$, $R = 12.6\,R_\odot$), we find that for critical-lobe overfilling $0.6 \lesssim e \lesssim 0.7$, depending on

© 1982 Nature Publishing Group

whether the primary is not rotating, rotating at the mean orbital angular velocity, or rotating at the orbital velocity at periastron. An eccentricity in this range is consistent with the observed $\dot{P}$ if the line of apsides makes an angle of 55°–65° with the line of sight (assuming the observation took place at periastron).

In conclusion, the observed P, $\dot{P}$ imply that A0538−66 is an eccentric binary system in which a neutron star with modest magnetic field ($B \lesssim 10^{11}$ G) approaches so close to its B star companion that massive accretion takes place. The region of hot gas giving rise to the optical outburst must be comparable in size with the separation of the two stars, and surrounding one or both of them. These deductions are based on data lasting only 4,000 s and observations of pulsations over longer periods would yield a great deal of information about this unique system. We urge that attempts be made to observe corresponding pulsations at optical and radio wavelengths.

We thank J. Halpern for assistance with the MPC analysis and F. R. Harnden for analysis of HRI data. J.E.G. is partially supported by an A. P. Sloan Fellowship and grant NAS8-30751.

Received 10 March; accepted 19 April 1982.

1. White, N. E. & Carpenter, G. F. *Mon. Not. R. astr. Soc.* **183**, 11–15P (1978).
2. Johnston, M. D. *et al. Astrophys. J. Lett.* **230**, L11–14 (1979).
3. Skinner, G. K. *Nature* **288**, 141–143 (1980).
4. Skinner, G. K. *et al. Astrophys. J.* **240**, 619–627 (1980).
5. Pakull, M. & Parmar, A. *Astr. Astrophys.* **102**, L1–4 (1981).
6. Charles, P. A. *et al. Space Sci. Rev.* **30**, 423–431 (1981).
7. Skinner, G. K. *Space Sci. Rev.* **30**, 441–446 (1981).
8. Brown, R. L. & Gould, R. J. *Phys. Rev.* **D1**, 2252–2256 (1970).
9. Schrier, E. J. *et al. Astrophys. J.* **204**, 539–547 (1976).
10. Sadeh, D. *et al. Astrophys. J.* (in the press).
11. Charles, P. A. *et al. Mon. Not. R. astr. Soc.* (submitted).
12. Davidson, K. & Ostriker, J. P. *Astrophys. J.* **179**, 585–598 (1973).
13. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296–316 (1979).
14. Lamb, F. K. *et al. Astrophys. J.* **184**, 271–283 (1973).
15. Bahcall, J. N. *A. Rev. Astr. Astrophys.* **16**, 241–264 (1978).
16. Avni, Y. *Astrophys. J.* **209**, 574–577 (1976).

Evidence for two pre-Flandrian palaeosols in Buchan, north-east Scotland

E. Rodger Connell*, Kevin J. Edwards† & Adrian M. Hall‡

* Department of Geography, University of Aberdeen, Aberdeen AB9 2UF, UK
† Department of Geography, University of Birmingham, Birmingham B15 2TT, UK
‡ Department of Geography, University of St Andrews, St Andrews KY17 9AL, UK

Evidence is presented for the existence of two distinct pre-Flandrian palaeosols at a single site, Kirkhill Quarry, in Buchan, north-east Scotland. This is the first such site reported from Scotland and it is possible that the two palaeosols represent interglacial episodes which are equivalent in age to last interglacial (compare with Ipswichian) and the penultimate interglacial (compare with Hoxnian)—a combination so far unknown in the British Isles. The stratigraphy also demonstrates that the part of the Buchan plateau under study may have been glaciated on at least three separate occasions.

Kirkhill Quarry (NGR NK 012 529) is located 13 km north-west of Peterhead, Aberdeenshire (Fig. 1*a*). The surface of the site lies at ~75 m OD and the bedrock consists of metasediments of the Dalradian Supergroup[1] intruded by a complex acid/basic dyke within which the quarry was opened.

The most complete development of the Pleistocene stratigraphy is seen in the quarry's east and north faces (Fig. 1*b*) and it is the evidence from these that is used to interpret the history of the site. Preservation of the sequence is a result of the protection afforded by the irregular bedrock surface consisting of very steep-sided channels or basins. The topographic location of these features and their association with glacigenic sediments suggest an origin by glacial or glacifluvial erosion facilitated, in part, by the weathered nature of the basic dyke rock.

The earliest phase of sedimentation at the site is recorded by gravels and coarse sands of fluvial (perhaps glacifluvial) origin. Typically this deposit has a light olive-brown colour (2.5 Y 5/4)[2] but the upper zones of both profiles consist of a 19-cm thick light-grey (10 YR 7/1) horizon overlying 7 cm of greyish-brown (10 YR 5/2) coarse sand containing ~1.4% organic matter. The latter is most clearly represented as a single organic band (~2 cm thick) in the east face. The north face stratigraphy is more complex with two main organic bands; the lower is ~1 cm thick and occurs within a blocky rubble matrix and the upper is ~6 cm thick and lies within bands of grey/yellow sands. There is deformation of the grey sand/organic sections of both profiles and a frost crack is visible in the east face. These horizons are overlain by coarse angular felsite debris which has probably been produced by frost shattering of the adjacent channel walls. This material is succeeded by till.

The till is present in both north and east faces where it is made conspicuous by its yellowish-brown matrix colour (10 YR 5/4) and its content of soft friable igneous and metamorphic clasts. In the north face section, the upper 70 cm of till have a strong brown colour (7.5 YR 5/6) with coarse distinct yellowish-red (5 YR 4/6–5/8) and greyish-brown/brown (10 YR 5/2–5/3) mottles. This would appear to be a weathering profile (discussed below) and it also displays signs of cryoturbation with the erection of clasts.

The till is buried beneath a head deposit which averages 60 cm in thickness and this is succeeded by further till deposits representing the last major event, or events, recognized. This situation is of some complexity because tills of both north-western (north face till derived from Moray Firth Mesozoic clays) and western (east face till derived from the igneous and metamorphic rocks of western Buchan) provenances are present. The availability of sections has not yet made it possible to determine which of these tills is the older. In contrast to the

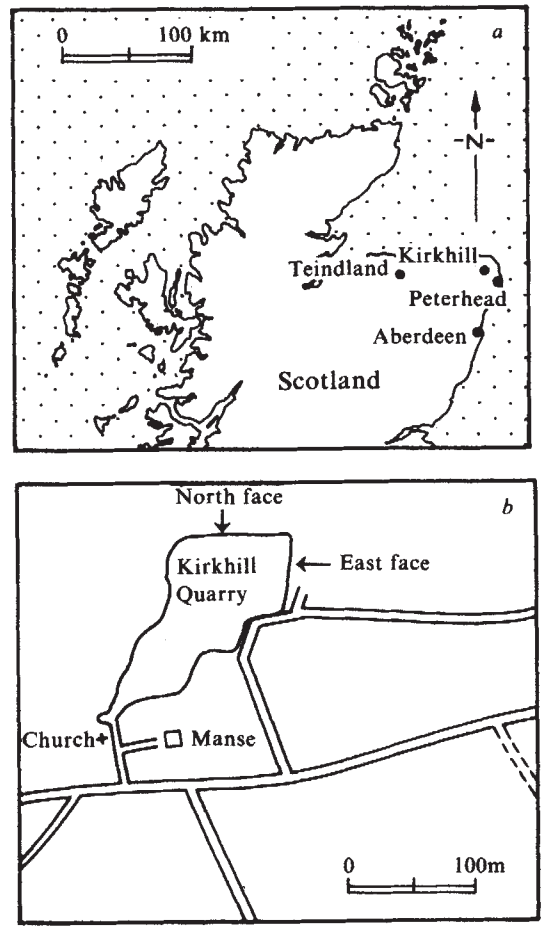


Fig. 1 *a*, Location of Kirkhill, Scotland. *b*, Kirkhill Quarry and location of sampling profiles.

0028–0836/82/240570—03$01.00
© 1982 Macmillan Journals Ltd

Table 1 Summary of pollen data

Local pollen assemblages	No. of sampling levels	
Upper palaeosol	North face	East face
KQ-6 *Alnus*	1	
Lower palaeosol		
KQ-5 *Calluna*		1
KQ-4 Gramineae—*Alnus*		2
KQ-3 Gramineae—*Alnus*–*Solidago*(+*Quercus*, *Picea*)	2	1
KQ-2 *Pinus*–*Alnus* (+*Tilia cordata*)		1
KQ-1 *Betula*—coryloid (+*Ulmus*, *Quercus*, *Picea*)		1

underlying and apparently severely weathered till, the upper tills have a fresh unweathered appearance and contain hard intact clasts.

The north face till is overlain by glacifluvial sand which fills a shallow channel. This sand has a humus iron podsol developed in it and appears to have been partially truncated by quarrying activity at some locations. Outside the channel limits, gleyed brown earths have developed in the upper tills.

Weathering in the light grey basal sands has resulted in the bleaching and softening of felsite clasts and the almost complete removal of such minerals as augite, hornblende and biotite from the fine sand fraction. Such a sequence of horizons resembles podsol-type soil profiles which have undergone weathering and translocation of iron and organic matter in a humid temperate climate.

The condition of the soft, friable clasts in the lower till would have precluded transportation and so a period of subaerial weathering after till deposition is indicated. In the north face weathered section of this till, clay mica has been altered to interstratified mica-smectite. X-ray diffraction analyses have identified the presence of lepidocrocite, an iron oxyhydroxide found in hydromorphic soils of temperate humid climate[3]. In thin section, mottles are seen as areas of iron oxide segregation and evidence of clay translocation is provided by numerous large angular papules of strongly orientated clay minerals, interpreted as void argillans[4] disrupted by subsequent cryoperturbation[5]. Taken together, these features indicate both gleying and translocation of clay, processes that are characteristic of B horizons of certain soil groups[6]. The upper 70 cm of the weathered till is thus interpreted as the truncated remnant of a soil formed under a humid temperate interglacial climate. The severity of clast weathering supports this interpretation.

Sediment samples from the basal and upper section of the north and east face profiles were prepared for pollen analysis by standard acetolysis and hydrofluoric acid procedures[7]. Sampling of organic layers for macroscopic plant and animal remains produced nothing. Only nine sampling levels produced sufficient pollen to permit useful discussion of the contained microfossils. Some of the spectra came from parallel sampling profiles and it was considered possible to place them in an internally consistent sequence by reference to lithostratigraphy and pollen content. The results are summarized in Table 1 and sample levels are indicated in Fig. 2. Pollen samples from the lower soil profile display some of the sharp inter-spectral changes found in the early buried podsol at nearby Teindland[8] and do not suggest the homogenization of assemblages which can result from microfaunal mixing or downflow of pollen through a permeable soil profile[9,10] (though such processes may still have occurred in part). The pollen assemblages have a mean arboreal representation (including shrubs) of 40.3% of total land pollen with a range of from 8.9% for *Calluna* assemblage KQ-5 to 96.1% for *Pinus-Alnus* assemblage KQ-2. The sequence KQ-1 to KQ-5 may well represent the relict registration of a vegetational landscape passing through an interglacial cycle[11,12] with a major local woodland component around the Kirkhill site. It is not possible to say whether the low amounts (1.0–5.0%) of such thermophilous taxa as *Ulmus* and *Quercus* in KQ-1 and KQ-3, *Tilia cordata* in KQ-2, as well as *Picea* in KQ-1 and KQ-3, indicate the local presence of elm, oak, lime and spruce or whether their pollen grains reached the site by long-distance wind transport from warmer areas to the south or east. If these taxa are of local origin (and not derived secondary pollen), their presence strongly supports an interglacial interpretation for the basal palaeosol.

A single sediment sample from the weathered section of the north face profile produced a pollen spectrum (KQ-6) dominated by *Alnus* (73.9% of the total land pollen) together with lower frequencies of *Betula* (4.3%) and coryloid (8.7%) grains. Such an assemblage would not be inconsistent with an interglacial phase.

Note that all pollen preparations contained microfragments of charcoal, while sieving for animal and plant remains also produced charcoal macrofragments in the basal organic levels. While evidence of burning is not limited to warm or temperate environments, it is more likely that such environments would provide a suitable context for the occurrence of natural fires.

Radiocarbon assay of three samples of organic material apparently representing both *in situ* soil (east face) and redeposited material (upper organic layer, north face) produced ^{14}C dates within the age range $33{,}810 \pm^{630}_{590}$ (SRR-1636) to $45{,}630 \pm^{1{,}740}_{1{,}430}$ (SRR-1635) yr b.p. (Fig. 2 and Table 2). The association of these ostensibly middle Devensian dates with pollen of a relatively thermophilous nature and linked with the stratigraphical context of this buried soil is considered anomalous. These dates must be regarded as minimal estimates of the true sample ages. This situation has been encountered with other Scottish interglacial sites and is generally thought to reflect sample contamination[13,14]. The $\delta^{13}C$ values lie within the normal range for terrestrial organic materials from European localities.

The sequence of events inferred from Kirkhill Quarry can be briefly summarized as: (1) erosion of channels/basins and deposition of fluvial or glacifluvial sands and gravels; (2) weathering and pedogenesis in humid temperate interglacial conditions to produce a podsol-like soil profile in the sands and gravels; (3) disturbance and burial of the podsolic profile by periglacial sediments followed by deposition of till; (4) weathering and soil development in the till under a second period of

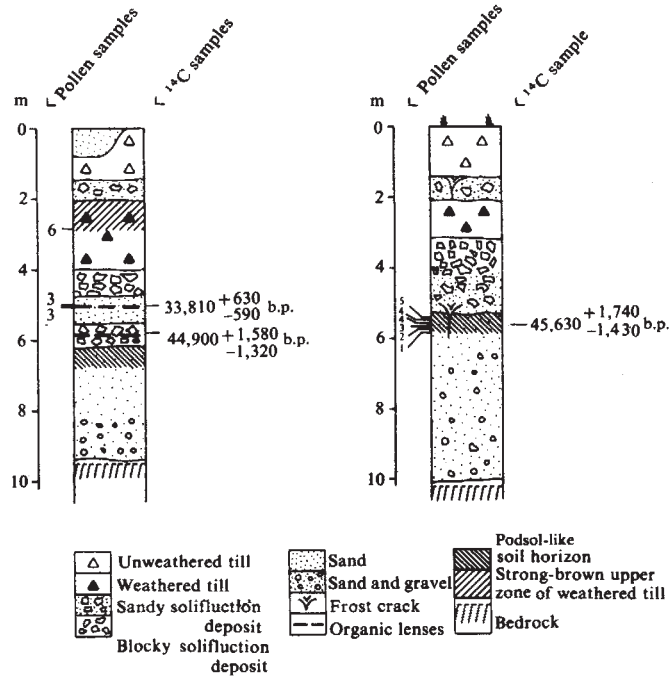


Fig. 2 Stratigraphic columns and positions of sampling levels for palynology (numbers indicate KQ–assemblage designations) and radiocarbon dating.

© 1982 Nature Publishing Group

Table 2 Radiocarbon dates from Kirkhill Quarry

Sample no.	Location	Depth below surface (m)	Date (yr b.p.)	$\delta^{13}C$ value
SRR-1635	East face	5.69–571	$45{,}630 \pm^{1{,}740}_{1{,}430}$	−26.3%
SRR-1636	North face	5.06–5.12	$33{,}810 \pm^{630}_{590}$	−27.0%
SRR-1637	North face	5.895–5.905	$44{,}900 \pm^{1{,}580}_{1{,}320}$	−25.7%

humid temperate climate; (5) periglacial disturbance of the soil profile and final glaciation of the site resulting in the deposition of the fresh, unweathered, upper tills; (6) erosion of channel and deposition of glacifluvial sands (north face); (7) Flandrian pedogenesis in tills and sands.

The Kirkhill Quarry sequence is important for several reasons. It is only the sixth polleniferous pre-late Devensian site known from Scotland and only the fourth claiming interglacial status[8,12,15–17]. The site is unique thus far in Scotland in providing evidence for two interglacial phases whose individual identity can be demonstrated in stratigraphical superposition within a single section. Kirkhill Quarry lies within the so-called 'Moraineless Buchan'[18,19] and outside the limits of the Moray Firth–Strathmore glaciation. As the latter is widely regarded as being of late Devensian age[20] it is possible that the upper tills at the site are even older. Despite this problem, the unweathered nature of these upper tills encourages the view that they date to a phase of Devensian glaciation. If this is the case then the simplest interpretation for the underlying interglacial horizons is that they are last interglacial (compare with Ipswichian) and penultimate interglacial (compare with Hoxnian) in age and the Kirkhill Quarry sequence would then be the first site for which this combination of deposits has been reported from the British Isles. The palaeofloras, whilst suggesting interglacial conditions, do not permit obvious correlation with the varied interglacial deposits from further afield[12,16,21–24]. The uppermost *Alnus*-dominated spectrum KQ-6 at Kirkhill may be compared with the *Alnus*–Gramineae–*Plantago lanceolata* pollen assemblages zone T-1 at nearby Teindland[8] which was assigned to the Ipswichian interglacial, and the *Alnus* zone ScB2 at Scandal Beck, Cumbria[25], which was tentatively referred to the Ipswichian partly by reference to the Teindland data. The Gramineae–*Alnus*–*Solidago* assemblage KQ-3 in the lower palaeosol at Kirkhill also resembles Teindland spectrum T-1, and assuming that two distinct pre-Flandrian interglacial phases are represented at Kirkhill then this may indicate the cyclical nature of interglacial vegetational change.

If the above reasoning is accepted, then the age of the upper palaeosol would be no younger than last interglacial. If it should prove to be older then the site becomes even more unusual in the Scottish context as the area possesses no known interglacial deposits earlier than the Hoxnian date claimed for Fugla Ness, Shetland[16]. If the upper palaeosol at Kirkhill is of Hoxnian age (and assuming no hiatuses), the basal podsol would chronologically be of Cromerian age and a profile combining both Hoxnian and Cromerian stage deposits is similarly unknown from the British Isles. If it is accepted that the two palaeosols represent distinct interglacial events and if the basal sands and gravels are glacifluvial in origin, then the stratigraphy demonstrates that this part of the Buchan plateau has been glaciated on at least three occasions. The earliest episode, indicated by the basal glacifluvial deposits, would be comparable in age with at least the Anglian glaciation of the English sequence[22].

We thank Banff and Buchan District Council for access to the site; Professor J. W. Parsons, Dr M. J. Wilson and Mr D. M. L. Duthie for chemical and mineralogical analyses; Dr D. D. Harkness for radiocarbon dates; Dr E. A. FitzPatrick for thin sections and advice; Dr P. C. Buckland who searched for animal remains; and Professor F. W. Shotton for comments on an earlier draft. E.R.C. acknowledges the receipt of a NERC research studentship.

Received 3 March; accepted 22 April 1982.

1. Harris, A. L. & Pitcher, W. S. in *A Correlation of Precambrian Rocks in the British Isles* (eds Harris, A. L. *et al.*) 52–75 (Geological Society of London, 1975).
2. *Munsell Soil Color Charts* (Munsell Color, Baltimore, 1975).
3. Schwertmann, U. & Taylor, R. M. in *Minerals in Soil Environments* (eds Dixon J. B. & Weed, S. B.) 145–180 (Soil Science Society of America, Madison, 1977).
4. Brewer, R. *Fabric and Mineral Analysis of Soils* (Krieger, New York, 1976).
5. Bullock, P. & Murphy, C. P. *Geoderma* **22**, 225–252 (1979).
6. Avery, B. W. *Soil Classification for England and Wales (Higher Categories)* (Soil Survey Technical Monogr. no. 14, Harpenden, 1980).
7. Faegri, K. & Iversen, J. *Textbook of Pollen Analysis* (Blackwell, Oxford, 1975).
8. Edwards, K. J. Caseldine, C. J. & Chester, D. K. *Nature* **264**, 742–744 (1976).
9. Dimbleby, G. W. *J. Soil Sci.* **12**, 1–11 (1961).
10. Havinga, A. J. *Geologie Mijnb.* **53**, 449–453 (1974).
11. Iversen, J. *Uppsala Univ. Årbok* **6**, 210–215 (1958).
12. Birks, H. J. B. & Peglar, S. M. *New Phytol.* **83**, 559–575 (1979).
13. Sissons, J. B. *Boreas* **10**, 1–17 (1981).
14. Shotton, F. W. *Geol. J.* **8**, 387–394 (1973).
15. FitzPatrick, E. A. *Nature* **207**, 621–622 (1965).
16. Birks, H. J. B. & Ransom, M. E. *New Phytol.* **68**, 777–796 (1969).
17. Edwards, K. J. & Connell, E. R. *Quat. Newslett.* **33**, 22–28 (1981).
18. Charlesworth, J. K. *Trans. R. Soc. Edinb.* **62**, 769–928 (1955).
19. Synge, F. M. *Scott. geogr. Mag.* **72**, 129–143 (1956).
20. Jardine, W. G. & Peacock, J. D. in *A Correlation of Quaternary Deposits in the British Isles* (eds Mitchell, G. F. *et al.*) 53–59 (Geological Society of London, 1973).
21. West, R. G. *New Phytol.* **85**, 571–622 (1980).
22. Mitchell, G. F., Penny, L. F., Shotton, F. W. & West, R. G. *A Correlation of Quaternary Deposits in the British Isles* (Geological Society of London, 1973).
23. Warren, W. P. *Geol. Surv. Ir. Bull.* **2**, 315–322 (1979).
24. Mangerud, J., Sønstegaard, E., Sejrup, H. P. & Haldorsen, S. *Boreas* **10**, 137–208 (1981).
25. Carter, P. A., Johnson, G. A. L. & Turner, J. *New Phytol.* **81**, 785–790 (1978).

Biological significance of surface flooding in warm-core ocean eddies

D. J. Tranter*, G. S. Leech* & D. J. Vaudrey†

* Division of Fisheries Research and
† Division of Oceanography, CSIRO Marine Laboratories,
PO Box 21, Cronulla, New South Wales 2230, Australia

Deep convective overturn has an important role[1] in enhancing the productivity of warm-core ocean eddies. An eddy (F) formed from the East Australian Current in early 1978 overwintered further to the south and developed a significant phytoplankton bloom the following spring. We describe here another process which tends to terminate such blooms by submerging winter water, relatively rich in nutrients, beneath poorer water of more recent origin.

Warm-core eddies shed by the Gulf Stream[2–4], Kuroshio[5–7] and East Australian Current[8–10] are islands of warm water interacting with a cooler atmosphere. In winter, the eddy surface cools and a deep mixed layer develops by convective overturn. This mixing process distributes nutrients uniformly throughout the isothermal layer (thermostad[11]), where they are available for photosynthesis; but if phytoplankton are mixed to depths as great as 200, 300 or even 400 m, their production is limited by the availability of light. With the onset of summer heating, a cap of warmer water forms, which isolates the winter water below and conserves the phytoplankton in the light. From that point onwards, the thermostad has a characteristic signature by means of which it can be tracked (G. R. Cresswell, work in preparation).

The properties of eddy J, which had overwintered north of Sydney in 1979, changed dramatically between September and October (Fig. 1). A warm surface cap (shaded) appeared in October with a nitrate concentration $<5\,\mu g\,N\,l^{-1}$ (compared with $20\,\mu g\,N\,l^{-1}$ the previous month), without appreciable increase in particulate organic matter ($5–15\,\mu g\,N\,l^{-1}$). There was no spring bloom like that which had occurred in eddy F the previous year[1]. Where had all the nitrate gone?

In September, eddy J had a neighbour to the north-east, eddy K, whose winter thermostad was warmer (by ~1 °C) and shallower (by ~100 m) than that of J[12]. Figure 2 shows the temperature–salinity characteristics of the surface mixed layer of J in August, September and October compared with the

0028–0836/82/240572—03$01.00
© 1982 Macmillan Journals Ltd

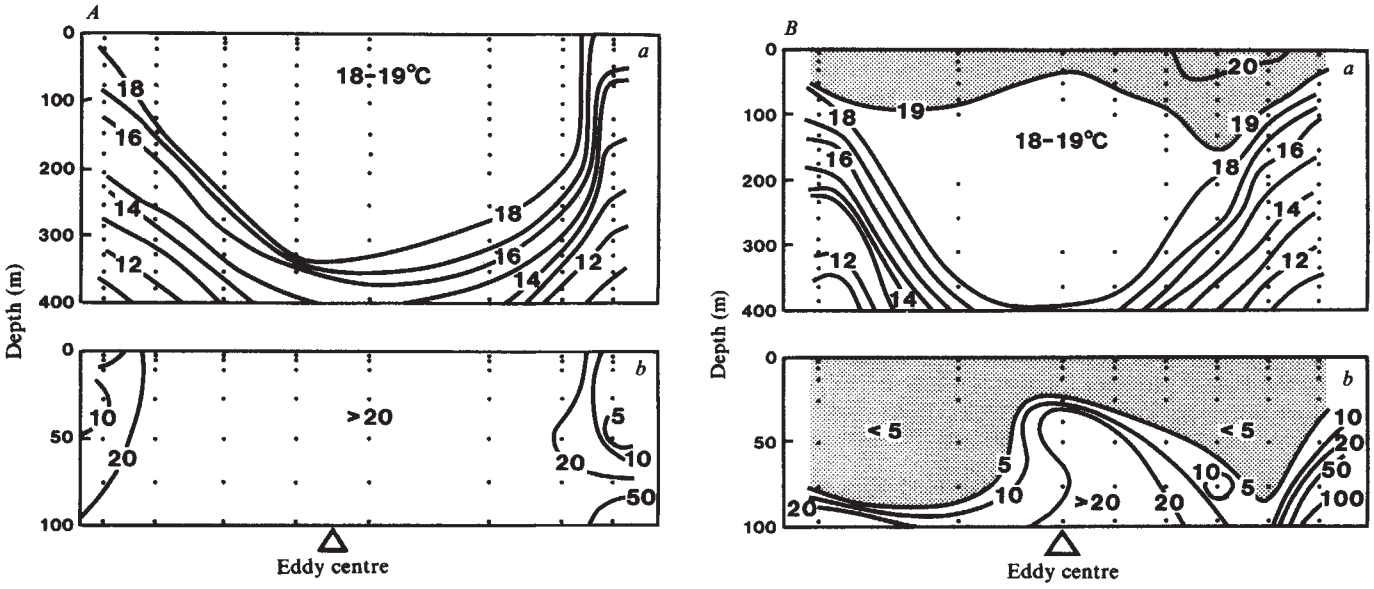


Fig. 1 Changes in the properties of eddy J between *A*, September and *B*, October 1979. *a*, Temperature (°C); *b*, inorganic nitrate (μg N l^{-1}). Eddy J developed a warm surface cap (shaded) and the nitrate concentration fell from >20 to <5 μg N l^{-1}.

September characteristics of K. For this purpose, the mixed layer was taken to be that stratum with a density range <0.1. The clusters of data points correspond to thermostads in the station profiles. Note that the J thermostad cooled by ~1 °C between August and September, an amount equivalent to the September differential between J and K. Note also that the October surface cap overlying the J thermostad had similar properties to the September thermostad of K. This warmer cap was evident in early October as an annulus around the outer edge of J and in late October as a surface blanket right across the eddy (G. R. Cresswell, work in preparation). We conclude that surface waters from eddy K spilled across the top of eddy J as an encircling ring which progressed towards its centre, eventually submerging it.

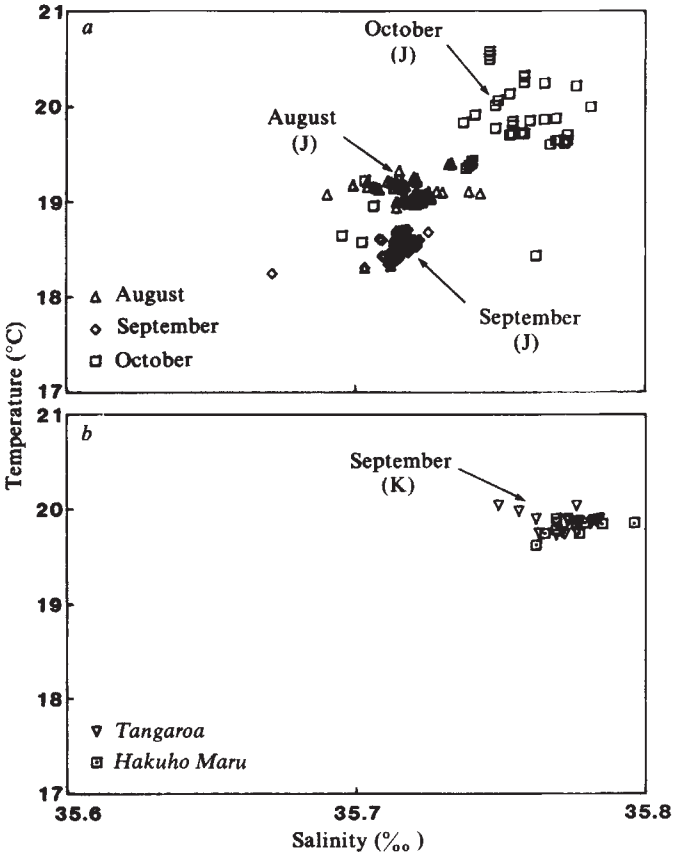


Fig. 2 Temperature–salinity characteristics of the surface mixed layer (where $\Delta\sigma t \not> 0.1$). *a*, J, August (SP9/79) September (SP10/79) and October (SP11/79). *b*, K, September (*Tangaroa* 1097, *Hakuho Maru* KH79-4). Note that the J thermostad cooled by ~1 °C before it was capped by a warm saline layer with similar properties to the K thermostad nearby.

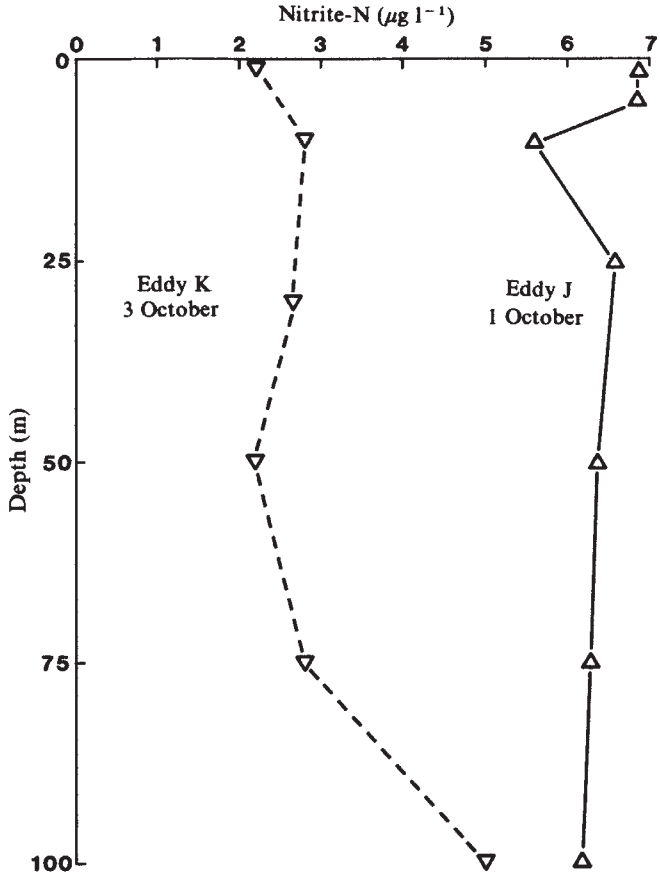


Fig. 3 Nitrite profiles through the K thermostad on 3 October (KH79-4 Stn 11 31°23′S 154°53.5′E); and the J thermostad on 1 October (SP10/79/238 33°20′S 152°46′E).

© 1982 Nature Publishing Group

Tomosada[7] observed the crest of the Kuroshio to over-ride adjacent eddies and submerge the winter thermostad. Eddy B was over-ridden by the East Australian Current some time between October 1977 and February 1978[10]. The relevance of submergence in the present context is in relation to its impact on the productivity of the system. Nutrients elevated in winter by convective overturn of the water column are submerged again by overflow and consequently are not utilized in primary production. The subsequent productivity of eddy J was therefore determined largely by the nutrients of the overflow from K. There are no September nitrate data available for K, but its nitrite concentration was lower than that of J (Fig. 3), perhaps because of shallower winter mixing. If Nilsson and Cresswell are correct in their assertion that most of the warm-core eddies shed by the East Australian Current are reabsorbed at a later date, then we suggest that significant eddy productivity may occur only among those few that move permanently beyond its reach.

Dr Basil Stanton of the New Zealand Oceanographic Institute made temperature–salinity data available to us from *Tangaroa* cruise 1097 and Dr N. Taga of the Ocean Research Institute provided nitrite data from *Hakuho Maru* cruise KH 79-4. We thank Drs J. S. Godfrey and G. R. Cresswell for helpful discussions.

Received 27 January; accepted 21 April 1982.

1. Tranter, D. J., Parker, R. R. & Cresswell, G. R. *Nature* **284**, 540–542 (1980).
2. Saunders, P. M. *Deep-Sea Res.* **18**, 1207–1219 (1971).
3. Richardson, P. L., Cheney, R. E. & Worthington, L. V. *J. geophys. Res.* **83**, 6136–6144 (1978).
4. Hulliwell, G. R. & Mooers, C. N. K. *J. geophys. Res.* **84**, 7707–7725 (1979).
5. Hata, K. *Oceanogr. Mag.* **21**, 13–29 (1969).
6. Cheney, R. E. *J. geophys. Res.* **82**, 5459–5468 (1977).
7. Tomosada, A. *Bull. Tokai reg. Fish. Res. Lab.* **94**, 59–102 (1978).
8. Andrews, J. C. & Scully-Power, P. *J. phys. Oceanogr.* **6**, 756–765 (1976).
9. Nilsson, C. S., Andrews, J. C. & Scully-Power, P. *J. phys. Oceanogr.* **7**, 659–669 (1977).
10. Nilsson, C. S. & Cresswell, G. R. *Prog. Oceanogr.* **9**, 133–183 (1980).
11. Seitz, R. C. *J. mar. Res.* **25**, 203 (1967).
12. Stanton, B. R. *N.Z. J. mar. Freshwat. Res.* **15**, 289–297 (1981).

New basal Namurian (Upper Carboniferous) fishes and crustaceans found near Glasgow

S. P. Wood

Hunterian Museum, University of Glasgow, Glasgow G12 8QQ, UK

I report here a recent discovery of new species of sharks, together with the most complete and best-preserved Carboniferous shark known, in Bearsden near Glasgow. Sharks of this age are significant because of their high diversification rate during the Carboniferous. This, coupled with their poor fossil record, has led to "almost as many classifications of Palaeozoic chondrichthyan fish as there are taxa in this group represented by adequate specimens"[1]. The new shark discoveries also throw considerable morphological light on relatively fragmentary material collected during the past few years from the Upper Mississippian (Lower Namurian) of Montana[2]. The fish found also include 11 new palaeoniscid (early ray-finned) fishes which represent the first known British Namurian marine species. The site has also yielded a crustacean assemblage which is probably the best preserved from Europe or even the Northern Hemisphere. The only comparable crustacean fauna, from beds of Wesphalian age at Mazon Creek in Illinois, is larger but is mainly preserved as moulds whereas the Bearsden material preserves the cuticle intact. The first (and last) major discovery of a substantial vertebrate and crustacean marine fauna from the British Carboniferous was made in 1879 when the Glencartholm site (Dumfries and Galloway) was found by Macconochie of the Scottish Geological Survey[3] in the Lower Carboniferous, Upper Border Group (Upper Dinantian, Asbian).

The Bearsden fossils are contained in a series of beds of marine shale which vary in lithology and have minor non-marine intercalations. Some specimens were collected from weathered exposures along the Manse Burn but most were obtained from a temporary quarry 5 × 3 m in area excavated in the south bank. This quarry is marked locality 1 on a detailed map at the NCC Geological Review Unit, Newbury to which professional enquiries should be directed. The *Crangopsis* marker band (Fig. 1) outcrops twice in the burn and is also included in this map together with the general dip and structure of beds A–E.

From an examination of the shelly fauna, R. B. Wilson (personal communication) places the Bearsden fauna at the horizon of the Top Hosie Limestone. Currie[4] placed the base of the Pendleian, the lowermost Namurian stage, just below this limestone.

Specimens collected during the excavation at Bearsden, which took place throughout summer 1981, have been deposited at the Hunterian Museum of the University of Glasgow (HM V) and the Royal Scottish Museum, Edinburgh (RSM GY).

Outside North America, Carboniferous marine sharks have only been described from Glencartholm near Langholm[5,6] whilst non-marine sharks are known exclusively from the Edinburgh area[6–10]. Both are of Asbian age in the Dinantian of Scotland[11]. In addition, xenacanths have been recorded from the Permo-Carboniferous of Europe[12].

None of the marine sharks discovered at Bearsden occurs at the Edinburgh or Glencartholm sites, and the non-marine

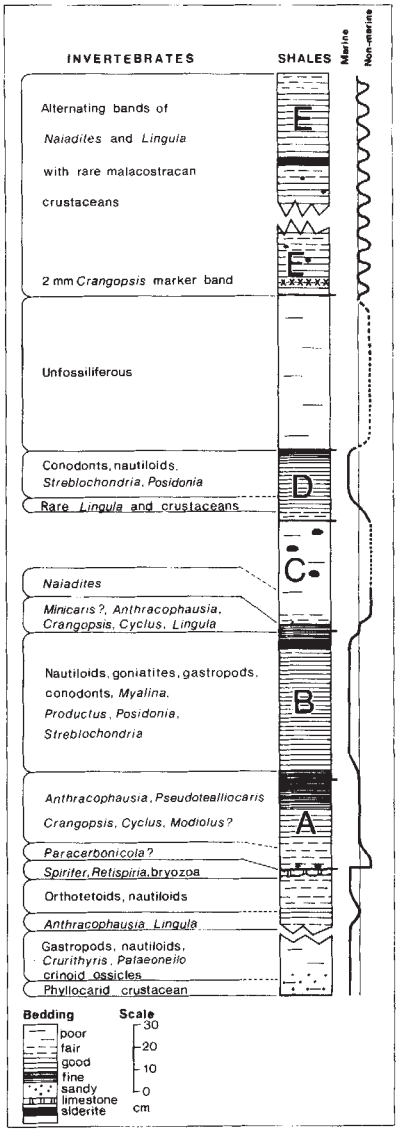


Fig. 1 Measured section in Manse Burn area. Lithological horizons used in the text are all approximate equivalents of Top Hosie limestone in Scotland.

0028–0836/82/240574—04$01.00

© 1982 Macmillan Journals Ltd

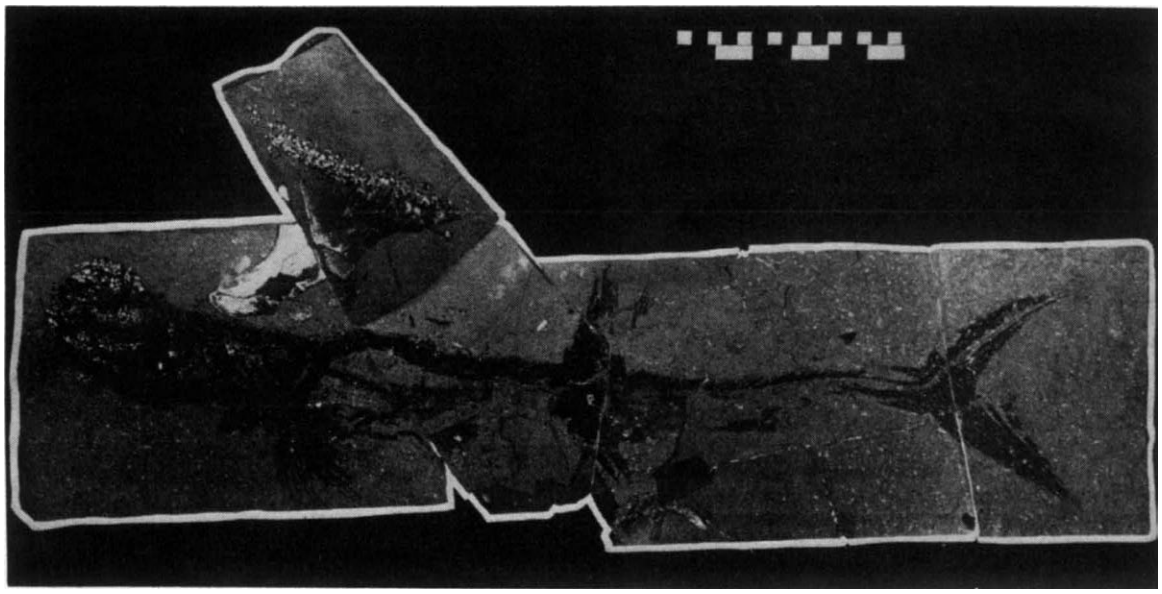


Fig. 2 Lateral view of superbly preserved anacanthous male shark (HM V8246). Photographed in methanol; upper scale bar in cm.

phalacanthous *Tristychius arcuatus* Agassiz is present at Bearsden and Edinburgh, but not at Glencartholm, despite many references to the contrary[7]. On the other hand, the only fish found at both, the phalacanthous *Onychoselache traquairi* Dick, is apparently absent at Bearsden. The Bearsden sharks are probably the best-preserved cartilaginous fish known from the Palaeozoic. One of them, HM V8246 (Fig. 2), should help to clarify the unsatisfactory systematics of anacanthous sharks. It could be *Cladodus neilsoni* Traquair, at present known only from the anterior end of a skeleton of similar age from East Kilbride and described by Traquair in the nineteenth century[13], or equally the Bear Gulch *Stethacanthus altonensis* (St John and Worthen)[1,14].

HM V8246 has a peculiar structure posteriorly attached to the dorsal spine near its base. It is well calcified, thick, apparently inflexible and triangular in shape. The posterio-dorsal margin is armoured with formidable monocuspid tooth-like denticles which increase in size posteriorly. Similar, though smaller, denticles form an anteriorly positioned 'helmet' on the head whilst behind the pelvic fins, external sexual organs identify the individual as male. If restored parallel to the body, the long axis posteriorly attached to each pectoral fin extends backwards well beyond the pelvic fins. These axes may have been used to grasp the female during copulation. The grotesque dorsal armament may also have had a sexual role as some extant forms, when excited, make extensive wounds in the flanks of females during foreplay.

S. altonensis and *C. neilsoni* may be synonymous as study of the type specimen of the latter (RSM 1911.62.52) suggests that the two posterior basal plate radials of the right pectoral fin are damaged, and it is the drifted distal portions lying against the axis cartilages which Traquair misinterpreted as axial radials. His interpretation has since been used by Moy-Thomas and Miles[15] and by Romer[16] when comparing reconstructed pectoral fins of other Palaeozoic sharks. The realization that the post-metapterygial axis of *C. neilsoni* is free of radials and that this condition may also be true for *S. altonensis* supports their being synonymous.

In RSM GY 1981.63.23, another Bearsden specimen of the same species, the dorsal spine and basal plate are exposed in cross-section in juxtaposition and the articulation of these elements is no different from that illustrated by Maisey[17] for the Upper Devonian shark *Cladoselache*. Maisey suggests that the unornamented dorsal spine of *Cladoselache* was deeply inserted. However, the undisturbed anatomy of HM V8246, a shark belonging to the same anacanthous group, contains an unornamented spine of similar design which is in 'life' position and is clearly shallow based. It may therefore be significant that shallow emplacement of the dorsal spine in *Cladoselache* was formerly suggested by Harris[18].

Additional Bearsden material including an isolated head (HM V8250) suggests an overall length of up to 1.3 m for *C. neilsoni* (=*S. altenensis?*). Lund (personal communication) considers HM V8250 to be female and he independently suspects synonomy between *Cladodus* and *Stethacanthus*, the latter being senior. Zangerl[19] noted the possibility of a generic relationship between the Westphalian *Symmorium* and *C. neilsoni*, and the newly discovered Bearsden material reinforces this suggestion, although the peculiar anterior dorsal structure is apparently absent in the former. Another important shark in the collection is only 12.5 cm long, has small cladodont teeth, lacks a fin spine and displays a series of distal radials above the vertebral column, suggesting the retention of a long-based dorsal fin as in *Chondrenchelys*. However, the chordacentra of the latter are absent. The poorly known bradyodont group is represented at Bearsden by a single complete suite of crushing teeth.

Eleven new palaeoniscid species seem to be represented. Many of the specimens are complete and all are well preserved. One aberrant form will require a new suprageneric taxon. This small fish has a large head, a raised denticle centrally positioned on each scale and long tapering projections extending beyond the pectoral and caudal fins, the former consisting of modified fulcral scales. The others collected include new species of *Elonichthys*, *Rhadinichthys* and *Mesopoma* (the latter identified by K. Lowney). In addition to the palaeoniscids, the deep-bodied chirodontid *Chirodus crassus* Traquair (Fig. 3) is present as several fine specimens which will allow accurate redescription, especially of the pelvic fins, hitherto presumed absent[20].

Acanthodians are represented by *Acanthodes sulcatus* Ag. The material includes a superb example (HM V8251), whilst in another specimen the first-known Carboniferous acanthodian braincase was recognized by A. L. Panchen and T. R. Smithson.

The sarcopterygians are poorly represented: dipnoans are so far absent, the only coelacanth seen is an isolated head, whilst the other crossopterygians occur as rare rhizodont and osteolepid fragments.

Another remarkable feature of the Bearsden discovery is the superbly preserved crustacean fauna, rivalling that from any British or European locality. It complements the well-known Westphalian fauna of Mazon Creek, Illinois, but is less diverse

© 1982 Nature Publishing Group

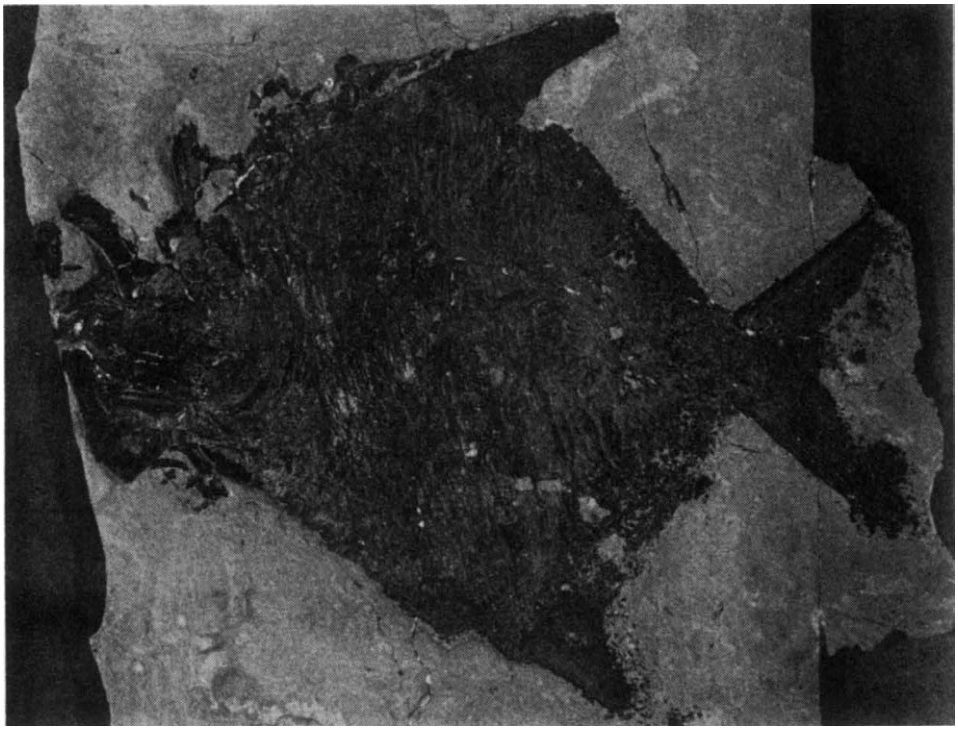


Fig. 3 *Chirodus crassus* (HM V8247), a deep-bodied palaeoniscoid 31 cm long, yielding new anatomical detail. Photographed in methanol.

than the North American fauna and consists of the malacostracans *Anthracophausia*, *Crangopsis*, *Pseudotealliocaris* and *Minicaris*?, together with the problematical '*Cyclus*'. This assemblage contains the near-shore and brackish water elements of the four marine communities proposed by Schram[21] which occur together, indicating a broader ecological niche than has been appreciated. Fine ornament and delicate structures are preserved in the cuticle of *Anthracophausia*, and W. D. I. Rolfe identified the raptorial appendages (Fig. 4) as the second antennae. This conclusion challenges the filter mode of feeding previously ascribed to *Anthracophausia*[22]. The occurrence of *Minicaris*? is noteworthy as this fossil is known from a single example in the Institute of Geological Sciences, Edinburgh, from a core sample of Asbian (middle Visean) age near Livingstone, Lothian Region[23].

Figure 1 shows the vertical distribution of the molluscs and some other elements of the fauna, and a relative salinity curve is drawn, using the invertebrates as indicators. Biofacies and lithofacies do not conveniently correspond, the former occasionally lagging behind the latter. The salinity profile through these overlap periods is therefore tentative. There are also constraints (except on *Acanthodes*) in fish distribution, some palaeoniscoids occurring only in the marine phases of B and D beds, whereas others are restricted to the transitional stratigraphy of basal B and upper A and so on. Similarly, shark restriction is apparent: *Cladodus* is only found in beds B and D and *Tristychius* in beds A (middle) and C, the latter probably representing Dick's[10] non-marine *Diplodoselache* fauna typical of the Oil Shales around Edinburgh. Further sampling should clarify this picture.

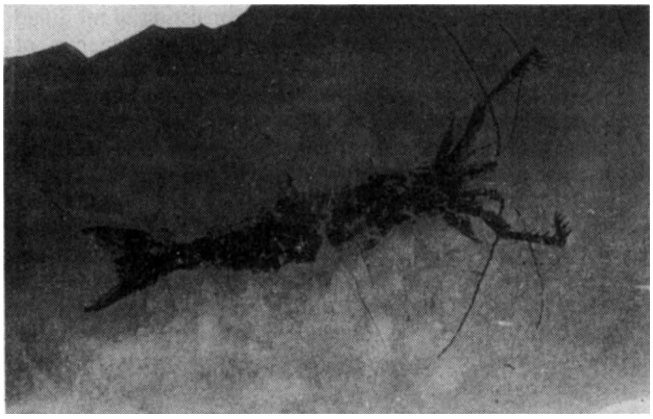


Fig. 4 *Anthracophausia dunsiana* (HM A2359) in dorsal view, an 8 cm long malacostracan crustacean. The anterior appendages were hitherto unknown in this animal.

The marine palaeoniscoid fauna at Bearsden is understandably different from the well-known non-marine species of the Scottish Midland Valley but as they also bear no specific relationship to the marine palaeoniscoids from Glencartholm, or the South African Witteberg assemblage[24], this raises interesting palaeoenvironmental questions.

In addition to the fish and crustaceans the collection contains microfauna and problematica and flora among which *Lepidophloios*, *Spathulopteris*, *Sphenopteris* and *Rhacopteris* have been recognized.

I thank the Leisure and Recreation Committee of Bearsden and Milngavie District Council for permission to quarry, Hunterian Museum staff for field assistance, Drs W. D. I. Rolfe, J. K. Ingham and A. L. Panchen for discussion, Dr R. B. Wilson for identifying and evaluating the molluscs, and Mr C. Cleal for identifying the flora. I acknowledge the help of Mr D. Barbour (drafting), Mr P. McGrail (typing) and Mr G. Grant (photography). I especially thank the Manpower Services Commission and Nature Conservancy Council for funding.

Received 16 February; accepted 31 March 1982.

1. Lund, R. *Ann. Carneg. Mus.* **45** (8), 161 (1974).
2. Melton, W. G. Jr *Proc. N. Am. Paleont. Conv.* **1**, 1202 (1971).
3. Traquair, R. H. *Trans. R. Soc. Edinb.* **30**, 15 (1881).
4. Currie, E. D. *Trans. R. Soc. Edinb.* **62**, 527 (1954).
5. Moy-Thomas, J. A. *Proc. zool. Soc. Lond.* **106**, 761 (1936).
6. Dick, J. R. F. & Maisey, J. G. *Palaeontology* **23**, 363 (1980).
7. Traquair, R. H. *Proc. Trans. R. Soc. Edinb.* **40**, 687 (1903).
8. Wood, S. P. *Scott. J. Geol.* **11**, 251 (1975).
9. Dick, J. R. F. *Trans. R. Soc. Edinb.* **70**, 63 (1978).
10. Dick, J. R. F. *Trans. R. Soc. Edinb.* **72**, 99 (1981).
11. George, T. N. *et al. Spec. Rep. geol. Soc. Lond.* **7**, 1 (1976).
12. Brongniart, C. *Bull. Soc. géol. Fr.* **3**, 16 (1888).
13. Traquair, R. H. *Trans. geol. Soc. Glasg.* **11**, 41 (1897).

© 1982 Nature Publishing Group

14. Williams, M. E. thesis, Univ. Kansas (1979).
15. Moy-Thomas, J. A. & Miles, R. S. *Palaeozoic Fishes* 2nd edn (Chapman & Hall, London, 1971).
16. Romer, A. S. *Vertebrate Paleontology* 3rd edn (University of Chicago Press, 1966).
17. Maisey, J. G. *Neues Jb. Geol. Paläont. Mh.* **1977**, 47 (1977).
18. Harris, J. E. *Scient. Publs. Cleveland Mus. nat. Hist.* **8**, 1 (1938).
19. Zangerl, R. in *Interrelationships of Fishes* (eds Greenwood, P. H., Miles, R. S. & Patterson, C.) 1 (Academic, London, 1973).
20. Traquair, R. H. *Trans. R. Soc. Edinb.* **29**, 343 (1879).
21. Schram, F. R. *J. Paleont.* **55**, 1 (1981).
22. Schram, F. R. in *Mazon Creek Fossils* (ed. Nitecki, M. H.) 159 (Academic, New York, 1979).
23. Schram, F. R. *Fieldiana Geol.* **40**, 1 (1979).
24. Gardiner, B. G. *Zool. J. Linn. Soc.* **48**, 423 (1969).

Carbon isotope ratios of apatite from fossil bone cannot be used to reconstruct diets of animals

Margaret J. Schoeninger* & Michael J. DeNiro†

Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

The reconstruction of animals' diets from measurements of stable isotope levels in fossils relies on the fact that the $^{13}C/^{12}C$ ratio of animal carbon reflects the $^{13}C/^{12}C$ ratio of dietary carbon[1,2]. Two phases in fresh bone, collagen and the carbonate occurring in apatite, the predominant bone mineral, have isotopic ratios that are related to the $^{13}C/^{12}C$ ratio of the diet[1]. The isotopic ratios of both phases have been used to study the diets of extant animals[3,4]. Reconstruction of the diets of fossil animals using the isotopic method has been limited to analysis of collagen preserved in bone[5–8]. It has not been possible to use the $^{13}C/^{12}C$ ratios of carbon in the inorganic phase of fossil bone for dietary reconstruction because most fossil bones contain significant amounts of calcium carbonate, deposited after the animal's death, that contribute to the CO_2 evolved from the bone during acid hydrolysis[3,4]. However, Sullivan and Krueger[9] recently presented data which led them to conclude that the $^{13}C/^{12}C$ ratio of CO_2 extracted by acid hydrolysis from the apatite phase of fossil bone records information about an animal's diet. We have now determined the $^{13}C/^{12}C$ ratios of both the apatite phase and the collagen of 24 fossil animal and human bones, and our results indicate that the $^{13}C/^{12}C$ ratio of fossil bone apatite cannot be used for dietary reconstruction.

Sullivan and Krueger[9] treated fossil bone with acetic acid to remove any secondary calcium carbonate that might have been present. They found that the $^{13}C/^{12}C$ ratios of the apatite carbonate fractions, which resisted acetic acid dissolution, were linearly correlated with $^{13}C/^{12}C$ ratios of the collagen for fossil as well as fresh bones. Because the collagen $^{13}C/^{12}C$ ratio is related to the isotopic composition of an animal's diet[1], they concluded that the $^{13}C/^{12}C$ ratios of apatite from fossil bone can be used to reconstruct aspects of the diets of animals that lived in the past.

There is a serious complication with the proposal[9] that the $^{13}C/^{12}C$ ratios of apatite in fossil bone can be used to reconstruct diet. Comparisons of the ^{14}C content of apatite with the ^{14}C content of collagen from the same bone or of other carbon-containing materials (such as charcoal) excavated in conjunction with the bone in question indicate that the apatite carbon can undergo exchange with carbon encountered in the post-mortem environment, either in groundwater or in the atmosphere (see, for example, refs 10–13). These diagenetic isotopic exchange processes can also affect the $^{13}C/^{12}C$ ratios of apatite[4]. To illustrate the magnitude of the stable isotopic shifts that can occur during diagenesis, we present the results of analysis of several suites of fossil bones.

We determined the $^{13}C/^{12}C$ ratios of the apatite and collagen of 6 modern and 24 fossil bones. These data, along with the ages and collection sites of the bones, are given in Table 1. Collagen was extracted from powdered bone as described previously[5] and its $^{13}C/^{12}C$ ratios determined following combustion by a modification of the Stump and Frazer method[14,15]. The $^{13}C/^{12}C$ ratios of the CO_2 liberated from apatite by reaction with anhydrous phosphoric acid were determined on separate aliquots of powdered bone that had been soaked in 50:50 (v:v) glacial acetic acid:water for two days before oxidation of the organic matter using sodium hypochlorite[1]. X-ray diffraction analysis demonstrated that the secondary calcium carbonate present in the fossil bones was removed by the acetic acid treatment.

We present a plot of the apatite $\delta^{13}C$ values against the collagen $\delta^{13}C$ values in Fig. 1. The line in Fig. 1 represents the relationship between collagen and apatite $\delta^{13}C$ values given by Sullivan and Krueger (see Fig. 1 legend). The data points for the six modern bones all lie close to this line. On the other hand, the differences between the observed apatite $\delta^{13}C$ value and that expected from the collagen $\delta^{13}C$ value for the fossil bones, based on the Sullivan-Krueger relationship, range from +4.6‰ to −12.1‰. This observation indicates that exchange processes have shifted the apatite $\delta^{13}C$ values from their original values, towards either more positive or more negative values, by as much as 12‰. Such a shift is significant when one considers that the average difference between the $\delta^{13}C$ values of C_3 and C_4 plants is about 14‰[16], while that between aquatic and most terrestrial food sources is about 7‰[1,8]. Determination of the

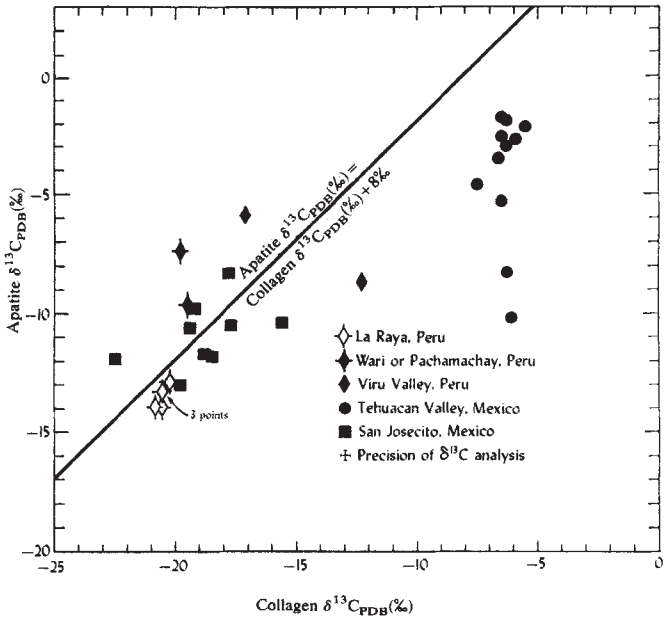


Fig. 1 The relationship between the $\delta^{13}C$ values of the apatite and collagen fractions of modern (open symbols) and fossil (closed symbols) bones analysed for this study. The line indicates the relationship given by Sullivan and Krueger[9]. Linear regression analysis of the data for fossil bones presented here produced the line: apatite $\delta^{13}C = 0.47$ collagen $\delta^{13}C - 1.29$‰ ($r = 0.80$). Sullivan and Krueger[9] indicated that apatite $\delta^{13}C$ = collagen $\delta^{13}C$ + 8‰. Linear regression analysis of their data yielded the line: apatite $\delta^{13}C = 1.07$ collagen $\delta^{13}C + 8.9$‰ ($r = 0.99$). The relationship given by Sullivan and Krueger is used in the discussion in the text. The values for $\delta^{13}C$ were calculated using the PDB belemnite carbonate as standard from the relationship:

$$\delta^{13}C = \left[\frac{(^{13}C/^{12}C)_{Sample}}{(^{13}C/^{12}C)_{Standard}} - 1\right] \times 1{,}000‰$$

* Permanent address: Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, Baltimore, Maryland 21209, USA.
† Also Program in Archaeology.

0028–0836/82/240577—02$01.00

© 1982 Macmillan Journals Ltd

Table 1 Collection location, taxonomic identity, age and $\delta^{13}C_{PDB}$ values for the apatite and collagen fractions of the modern and fossil bones analysed for this study

Location	Species	Age (yr BP)	Apatite $\delta^{13}C_{PDB}$ (‰)	Collagen $\delta^{13}C_{PDB}$ (‰)
La Raya, Peru	*Lama pacos*	0	−13.3	−20.5
	Lama pacos	0	−12.9	−20.2
	Lama pacos	0	−14.0	−20.6
	Lama pacos	0	−14.0	−20.8
	Lama pacos	0	−13.3	−20.5
	Lama pacos	0	−13.3	−20.5
Pachamachay, Peru	*Lama glama*	4,500–3,750	−9.7	−19.5
San Josecito, Mexico	*Canis diris*	~20,000	−10.4	−15.6
	Cervus sp.	~20,000	−10.6	−19.4
	Cervus sp.	~20,000	−9.9	−19.2
	Equus sp.	~20,000	−11.9	−22.7
	Equus sp.	~20,000	−13.0	−19.8
	Felis concolor	~20,000	−11.8	−18.5
	Felis concolor	~20,000	−11.7	−18.8
	Ursus arctos	~20,000	−10.5	−17.7
	Ursus arctos	~20,000	−8.3	−17.8
Tehuacan Valley, Mexico	*Homo sapiens*	1,300–500	−8.3	−6.3
	Homo sapiens	1,300–500	−2.2	−5.5
	Homo sapiens	1,300–500	−2.7	−5.9
	Homo sapiens	1,300–500	−1.9	−6.3
	Homo sapiens	1,300–500	−3.5	−6.6
	Homo sapiens	1,300–500	−3.0	−6.3
	Homo sapiens	1,300–500	−10.2	−6.1
	Homo sapiens	1,300–500	−5.3	−6.5
	Homo sapiens	2,800–2,150	−1.8	−6.5
	Homo sapiens	2,800–2,150	−4.6	−7.5
	Homo sapiens	7,000–5,400	−2.6	−6.5
Viru Valley, Peru	*Homo sapiens*	2,200–1,500	−8.7	−12.3
	Homo sapiens	2,200–1,500	−5.9	−17.1
Wari, Peru	*Lama glama*	1,350–1,160	−7.4	−19.8

The ages given for the bones are generally based on dates obtained for other materials excavated with them and hence provide minimum and maximum limits on the actual ages.

relative amounts of C_3 versus C_4 plants or of aquatic versus terrestrial foods in the diet are the two applications of the isotopic method of dietary analysis that have been made to date[5–8]. Uncertainties in the estimates of diet $\delta^{13}C$ values on the order of 5–10‰, resulting from diagenetic exchange processes, would thus make dietary reconstruction based on fossil bone apatite $\delta^{13}C$ values completely unreliable.

Our treatment of the data reported here, as well as that used by Sullivan and Krueger[9], is based on the assumption that the fossil bone collagen $\delta^{13}C$ values have not been altered by diagenetic processes. This may not be a valid assumption[3–5]. However, some of the apatite $\delta^{13}C$ values presented in Table 1 can be interpreted without resorting to comparisons with the corresponding collagen $\delta^{13}C$ values. Eight of the 11 samples from the Tehuacan Valley come from a single period of the occupation of the site[5]. There is no evidence to suggest differences in status among these individuals[17,18]. Accordingly, it is reasonable to expect that their diets, and consequently the bone isotope ratios, would be similar. The $\delta^{13}C$ values of collagen from these samples agree to within 1‰, which is consistent with this expectation. The $\delta^{15}N$ values of collagen, which have also been shown to be determined by diet, for these same eight samples have a range of only 1‰[5]. Note, however, that the apatite $\delta^{13}C$ values for these samples range from −1.9‰ to −10.2‰. These observations indicate that the apatite carbon in these bones has undergone post-mortem exchange. The large range of $\delta^{13}C$ values suggests that individual bones have exchanged with carbon sources of different $^{13}C/^{12}C$ ratios and/or have undergone different amounts of exchange.

The results presented here demonstrate that bone apatite carbonate undergoes exchange with carbon encountered after the death of the animal and that these diagenetic exchange processes can shift the apatite $\delta^{13}C$ values by at least as much as 12‰. Uncertainties of this magnitude in estimates of diet $\delta^{13}C$ values cannot be tolerated, since differences in the $\delta^{13}C$ values of different foodstuffs whose relative consumption is of interest are only 7–14‰. The apatite carbon in some fossil bones may not have undergone post-mortem exchange. However, until a method is developed to identify such isotopically unaltered apatite, we conclude that $\delta^{13}C$ values of fossil bone apatite cannot be used to reconstruct ancient diets.

The following individuals supplied or helped us obtain fossil bones: M. West, J. E. Ericson, C. C. Patterson, E. Wing, R. S. MacNeish, J. Wheeler, A. Romano and D. P. Whistler. L. Armstrong and D. Winter performed the X-ray diffraction analyses and/or isotopic measurements. This work was supported by NSF grant BNS 79-24756.

Received 27 November 1981; accepted 2 April 1982.

1. DeNiro, M. J. & Epstein, S. *Geochim. cosmochim. Acta* **42**, 495–506 (1978).
2. Teeri, J. A. & Schoeller, D. A. *Oecologia* **39**, 197–200 (1979).
3. DeNiro, M. J. & Epstein, S. *Science* **201**, 906–908 (1978).
4. Land, L. S., Lundelius, E. L. Jr & Valastro, J., Jr *Palaeogeogr., Paleoclimatol., Paleoecol.* **32**, 143–151 (1980).
5. DeNiro, M. J. & Epstein, S. *Geochim. cosmochim. Acta* **45**, 341–351 (1981).
6. van der Merwe, N. J. & Vogel, J. C. *Nature* **276**, 815–816 (1978).
7. Burleigh, R. & Brothwell, D. *J. Archaeol. Sci.* **5**, 355–362 (1978).
8. Tauber, H. *Nature* **292**, 332–333 (1981).
9. Sullivan, C. H. & Krueger, H. W. *Nature* **292**, 333–335 (1981).
10. Hassan, A. A., Termine, J. D. & Haynes, C. V. Jr *Radiocarbon* **19**, 364–374 (1977).
11. Haas, H. & Banewicz, J. *Radiocarbon* **22**, 537–544 (1980).
12. Haynes, V. *Science* **161**, 687–688 (1968).
13. Tamers, M. A. & Pearson, F. J. Jr *Nature* **208**, 1053–1055 (1965).
14. Stump, R. K. & Frazer, J. W. *Nuclear Sci. Abs.* **28**, 746 (1973).
15. Northfelt, D. W., DeNiro, M. J. & Epstein, S. *Geochim. cosmochim. Acta* **45**, 1895–1898 (1981).
16. Smith, B. N. & Epstein, S. *Pl. Physiol.* **47**, 380–384 (1971).
17. Anderson, J. E. in *The Prehistory of the Tehuacan Valley* (ed. Byers, D. S.) Vol. 1, 91–113 (University of Texas Press, Austin, 1967).
18. MacNeish, R. S. in *The Prehistory of the Tehuacan Valley* (ed. Byers, D. S.) Vol. 1, 290–309 (University of Texas Press, Austin, 1967).

©1982Nature Publishing Group

Oxygen production by endosymbiotic algae controls superoxide dismutase activity in their animal host

James A. Dykens & J. Malcolm Shick

Department of Zoology, University of Maine, Orono, Maine 04469, USA

Aerobic and aerotolerant organisms have evolved defenses against the toxic effects of molecular oxygen[1]. One protective mechanism involves the breakdown of the harmful superoxide radical by the enzyme superoxide dismutase (SOD). However, levels of oxygen elevated only slightly above normal atmospheric (P_{O_2} of 159 mm Hg) may overpower a cell's defense systems[2–4]. Although most animals do not naturally encounter oxygen pressures above 1 atm, hyperbaric oxygen levels normally occur in the tissues of marine animals that harbour intracellular algal symbionts, which in light generate more oxygen than is consumed by the combined host and symbionts[5–10]. We have now found that sea anemones *Anthopleura elegantissima* (Brandt) containing zooxanthellae (the symbiotic dinoflagellate *Symbiodinium microadriaticum*) in their gastrodermal tissues have SOD activities nearly two orders of magnitude greater than individuals totally lacking zooxanthellae.

Using oxygen microelectrodes we have recorded P_{O_2} values of up to 328 mm Hg in *A. elegantissima* gastrodermal tissues when the anemone is exposed to light intensities comparable with 33% of sunlight (Fig. 1). The bubbles that form on the surface of the anemones in bright light are likely to be composed of pure oxygen, as is the case in reef corals[7]. Thus, the host is subjected to a continuous flux of photosynthetically produced hyperbaric oxygen. Much attention has been paid to the mutualistic aspects of alga–invertebrate symbioses (such as transfer of photosynthetically fixed carbon compounds from algae to host in exchange for organic and inorganic metabolities)[11,12], yet little consideration has so far been given to the possibility of oxygen toxicity.

Molecular oxygen preferentially undergoes univalent reductions, one product of which is the highly reactive superoxide anion(O_2^-)[13]. Whether it is the superoxide anion *per se* or the hydroxyl radical (which may be produced by a reaction between O_2^- and H_2O_2) that is so destructive to cells remains debatable[14,15]. It is known, however, that O_2^- induces lipid peroxidation, depolymerizes hyaluronate, inactivates ribonuclease and destroys phagocytized bacteria when produced by activated polymorphonuclear leucocytes[16,17]. Virtually all aerobic organisms examined contain at least one form of the enzyme SOD[18] which catalyses the reaction $O_2^- + O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$, and so prevents the cellular accumulation of O_2^- and its resultant damage.

We considered that *A. elegantissima* symbiotically associated with greater numbers of zooxanthellae would have proportionally higher levels of SOD activity, to cope with the higher rate of oxygen production. This has been confirmed: SOD activity was found to be positively correlated with chlorophyll content in symbiotic anemones (Fig. 2).

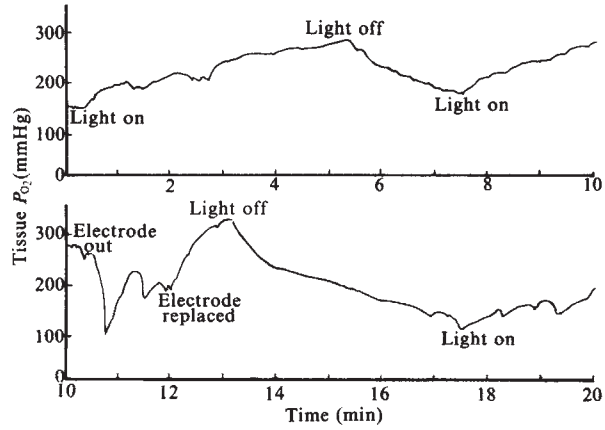


Fig. 1 Continuous record of changes in P_{O_2} in gastrodermal tissue of the zooxanthella-containing sea anemone *A. elegantissima* during illumination. Using a micromanipulator, a glass oxygen microelectrode (Transidyne General) was inserted into intact anemones at the base of the tentacles and illumination (315 μEinstein m^{-2} s^{-1}) was provided by high-intensity fibre optics. A few crystals of menthol were added to the seawater to reduce the anemone's movements so that the electrode would remain in position.

Anemones collected from the upper areas of the intertidal zone contain significantly ($P < 0.01$) more chlorophyll than anemones collected from the lower intertidal ($\bar{x}$ = 5.21 μg per mg protein, s.d. = 2.28, $n = 12$ compared with 2.67 μg per mg protein, s.d. = 1.07, $n = 17$) (Fig. 2). Photosynthetically produced carbohydrates are an important source of nutrition in *A. elegantissima*[19,20]. As compensation for reduced feeding time, high intertidal anemones may rely more heavily on algal photosynthate than low intertidal individuals and would, therefore, maintain a greater photosynthetic capacity. Regardless, SOD activity per μg chlorophyll is identical in high intertidal ($\bar{x}$ = 2.42 U per μg chlorophyll, s.d. = 1.89, $n = 12$) and low intertidal anemones ($\bar{x}$ = 2.42 U per μg chlorophyll, s.d. = 1.74, $n = 17$), supporting the idea that SOD activity of the host is modulated by the amount of photosynthetically generated O_2 and O_2^- in the tissues. It has also been noted that *Euglena gracilis* raised phototrophically has almost three times the SOD activity of *E. gracilis* raised heterotrophically[21].

To examine further the relationship between photosynthetic activity and sea anemone SOD activity, we divided members of a single clone of high intertidal *A. elegantissima* into four groups, which were given different treatments as follows: (1) sunlight (control); (2) sunlight plus 10^{-5} M 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) which inhibits photosynthesis; (3) darkness; and (4) darkness, plus constant exposure to 3 atm O_2. As expected, exposure to exogenous hyperbaric O_2 caused an increase in SOD activity of 62% (Table 1). Chlorophyll content tended to decrease (although not significantly) in both dark groups and in the DCMU-treated anemones (Table 1), in agreement with previous reports on the loss of zooxanthellae in both anemones and corals deprived of light[20,22]. The loss of photosynthetically produced O_2 by either

Table 1 Superoxide dismutase activity and chlorophyll content in *A. elegantissima*

	Sunlight (control)	Sunlight + 10^{-5} M DCMU	Darkness	Darkness + 3 atm O_2
SOD/protein (U mg^{-1})	3.81 ± 1.84*	1.41 ± 0.47†	2.90 ± 0.81*†	6.18 ± 1.71‡
Chl/protein (μg mg^{-1})	0.8423 ± 0.4630*	0.5424 ± 0.1452*	0.5406 ± 0.1961*	0.5233 ± 0.1662*

SOD activity and chlorophyll (Chl) content measured after 14 days in each of four treatments: 15 °C, $n = 10$ in each group, mean ± s.d. Means not significantly different by Tukey's test have same symbols. Control and DCMU-treated anemones were held outdoors in glass aquaria and exposed to at least 6 h direct sunlight (~860 μE m^{-2} s^{-1}) plus an additional 4–5 h of indirect sunlight (~680 μE m^{-2} s^{-1}) each day. Temperature (15 °C) was maintained with a thermostatted immersion refrigerator. Both dark-treated groups were held in a constant environment room at 15 °C.

0028–0836/82/240579—02$01.00

© 1982 Macmillan Journals Ltd

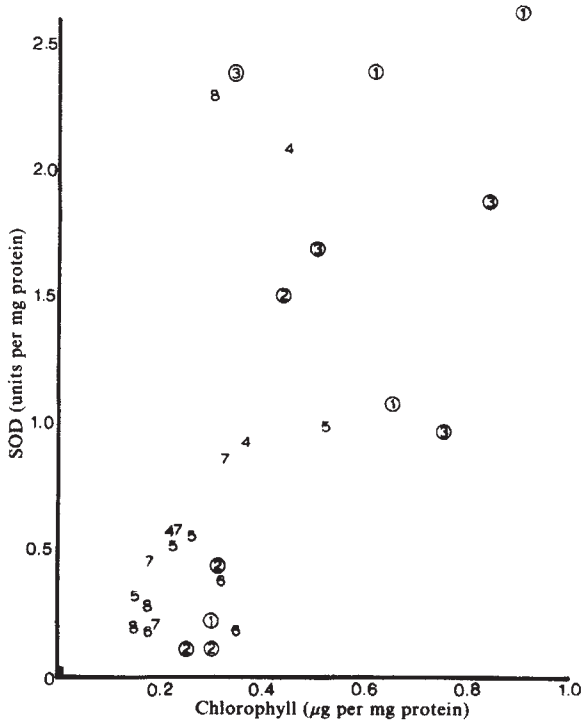


Fig. 2 Correlation between SOD activity and total chlorophyll in individuals from eight different clones of the sea anemone *A. elegantissima*. Clone-mates share numbers and symbiotic anemones collected from the upper areas of the intertidal zone are circled. The mean SOD activity of six aposymbiotic anemones totally lacking chlorophyll is represented by the single solid square (mean SOD activity $= 1.14 \times 10^{-2}$ U per mg animal protein). Regression: $y = 2.575x - 0.013$, $r = 0.738$, $P < 0.001$, $n = 35$. SOD activity, expressed in Units according to McCord and Fridovich[27], was determined using a Clark-type polarographic oxygen electrode and a procedure modified from Tyler[28], and from Marshall and Worsfold[29]. After homogenizing the anemone, the algal cells were removed by centrifugation (3,000*g*, 15 min) and the SOD activity reported is that of the resulting animal supernatant. Chlorophylls *a* and c_2 were calculated using the equations of Jeffrey and Humphrey[30], following two 10 h extractions in 100% acetone. No chlorophyll could be detected in the homogenization supernatant using this spectrophotometric procedure or by using the more sensitive fluorometric technique of Yentsch and Menzell[31], thus verifying that the algal cells had not been disrupted by the homogenization procedure and that the SOD activity is that of the animal tissue. Protein was determined according to Itzhaki and Gill[32], and, after tissue digestion in 5% KOH, by the technique of Bradford[33].

chemical inhibition or darkness decreased host SOD activity. It therefore appears that host SOD activity is altered in response to the amount of O_2 generated photosynthetically by its intracellular algal symbionts.

The net production of oxygen by alga–invertebrate associations has been interpreted almost exclusively as an indication of the nutritional contribution of the algae to the host[11,12,23], although intracellularly produced oxygen has been implicated as a buffer against environmental hypoxia and as a factor modifying behaviour in sea anemones[24–26]. The high levels of photosynthetically generated oxygen are a potentially deleterious factor in these symbioses. To exploit the benefits of the association while avoiding oxygen toxicity, the host sea anemones must maintain levels of SOD in direct proportion to their algal complement.

We thank Dr Bruce Sidell for technical advice, discussions and critical reading of the manuscript, Dr Leonard Muscatine for comments on the manuscript, and William Zamer for collecting and maintaining the anemones and for helpful discussions. This research was supported by the University of Maine Faculty Research Fund and by NSF grant PCM 79-11027 (Regulatory Biology) to J.M.S.

Received 18 March; accepted 20 April 1982.

1. Halliwell, B. *Cell Biol. int. Rep.* **2**, 113–128 (1978).
2. Fridovich I. *Science* **27**, 462–466 (1977).
3. Yagi, K. & Ohishi, N. in *Biochemical and Medical Aspects of Active Oxygen* (eds Hayaishi, O. & Asada, K.) 299–307 (University of Tokyo Press, Tokyo, 1977).
4. Barthelemy, L., Beland, A. & Chastel, C. *Resp. Physiol.* **44**, 261–268 (1978).
5. Pearse, V. B. *Biol. Bull.* **147**, 641–651 (1974).
6. Wethey, D. S. & Porter, J. W. in *Coelenterate Ecology and Behavior* (ed. Mackie, G. O.) 59–66 (Plenum, New York, 1976).
7. Crossland, C. J. & Barnes, D. J. *Mar. Biol.* **40**, 185–194 (1977).
8. Svoboda, A. & Porrmann, T. in *Nutrition in the Lower Metazoa* (eds Smith, D. C. & Tiffon, Y.) 87–99 (Pergamon, New York, 1980).
9. Trench, R. K., Wethey, D. S. & Porter, J. W. *Biol. Bull.* **161**, 180–198 (1981).
10. Mangum, C. P. & Johansen, K. *Pacific Sci.* (in the press).
11. Trench, R. K. *A. Rev. Pl. Physiol.* **30**, 485–531 (1979).
12. Muscatine, L. in *Primary Productivity in the Sea* (ed. Falkowski, P. G.) 381–402 (Plenum, New York, 1980).
13. Hill, H. A. O. *Ciba Fdn Symp.* **65**, 5–17 (1979).
14. Halliwell, B. in *Superoxide and Superoxide Dismutases* (eds Michelson, A. M., McCord, J. M. & Fridovich, I.) 335–349 (Academic, New York, 1977).
15. Fee, J. A. in *Biological and Clinical Aspects of Superoxide and Superoxide Dismutase* (eds Bannister, W. H. & Bannister, J. V.) 41–48 (Elsevier, Amsterdam, 1980).
16. Fridovich, I. *Science* **201**, 875–880 (1978).
17. Roos, D. & Weening, R. S. *Ciba Fdn Symp.* **65**, 225–262, (1979).
18. Hassan, H. M. & Fridovich, I. in *Enzymic Basis of Detoxication* (ed. Jakoby, W. B.) 311–332 (Academic, New York, 1980).
19. Trench, R. K. *Proc. R. Soc.* B**177**, 237–250 (1971).
20. Fitt, W. K. & Pardy, R. L. *Mar. Biol.* **61**, 199–205 (1981).
21. Asada, K., Kanematsu, S. & Uchida, K. *Archs Biochem. Biophys.* **179**, 243–256 (1977).
22. Steele, R. D. *J. Zool. Lond.* **179**, 387–405 (1976).
23. Muscatine, L., McCloskey, L. R. & Morian, R. E. *Limnol. Oceanogr.* **26**, 601–611 (1981).
24. Pearse, V. B. *Biol. Bull.* **147**, 630–640 (1974).
25. Fredericks, C. A. *Mar. Biol.* **38**, 25–28 (1976).
26. Shick, J. M. & Brown, W. I. *J. exp. Zool.* **201**, 149–155 (1977).
27. Fridovich, I. & McCord, J. *J. biol. Chem.* **244**, 6049–6055 (1969).
28. Tyler, D. D. *Biochem. J.* **147**, 493–504 (1975).
29. Marshall, M. J. & Worsfold, M. *Analyt. Biochem.* **86**, 561–573 (1978).
30. Jeffrey, S. W. & Humphrey, G. H. *Biochem. Physiol. Pflanzen.* **167**, 191–194 (1975).
31. Yentsch, C. S. & Menzel, D. W. *Deep-Sea Res.* **10**, 221–231 (1963).
32. Itzhaki, R. F. & Gill, D. M. *Analyt. Biochem.* **9**, 401–410 (1964).
33. Bradford, M. M. *Analyt. Biochem.* **72**, 248–254 (1976).

Central connections of the retinal ON and OFF pathways

Peter H. Schiller

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Slaughter and Miller have recently demonstrated that in the isolated eye cup of the mudpuppy and the rabbit, DL-2-amino-4-phosphonobutyric acid (APB) reversibly blocks the ON responses in the retina: on infusion of APB, ON bipolar, ON amacrine and ON ganglion cells become unresponsive; receptors, horizontal cells, OFF bipolars, OFF amacrines and OFF ganglion cells are unaffected[1]. The centre-surround organization of OFF-centre ganglion cells appears unaltered. I have used these findings to determine how the ON-centre and OFF-centre retinal ganglion cells, whose characteristics have been extensively studied in mammals[2–7], exert their influence on the neurones of the central visual system. Recordings made in the lateral geniculate nucleus (LGN) and striate cortex of the rhesus monkey while the retinal ON responses were reversibly blocked with APB showed that in the LGN, retinal APB infusion blocked the responses of ON-centre cells and had little effect on OFF-centre cells. The centre-surround organization of OFF-centre cells was unaffected. APB infusion eliminated the light-edge responses of cortical cells, revealed a dark-edge response in some, but had no discernible effect on orientation and direction specificities. These results suggest that the ON and OFF systems do not interact significantly at the level of the LGN, but do so in the striate cortex.

Surgery was performed on rhesus monkeys under sodium pentobarbital anaesthesia. During recording, animals were paralysed with Flaxedil and were respirated with a 75–25% mixture of NO_2–O_2. The eye was perfused through a tube inserted behind the limbus into the vitreous chamber. The perfusate left the eye through a second tube. The control perfusate, which was similar to that reported by Cunningham and Miller[8], was dripped at the rate of 2–4 ml min^{-1}. A stopcock

0028–0836/82/240580—04$01.00

© 1982 Macmillan Journals Ltd

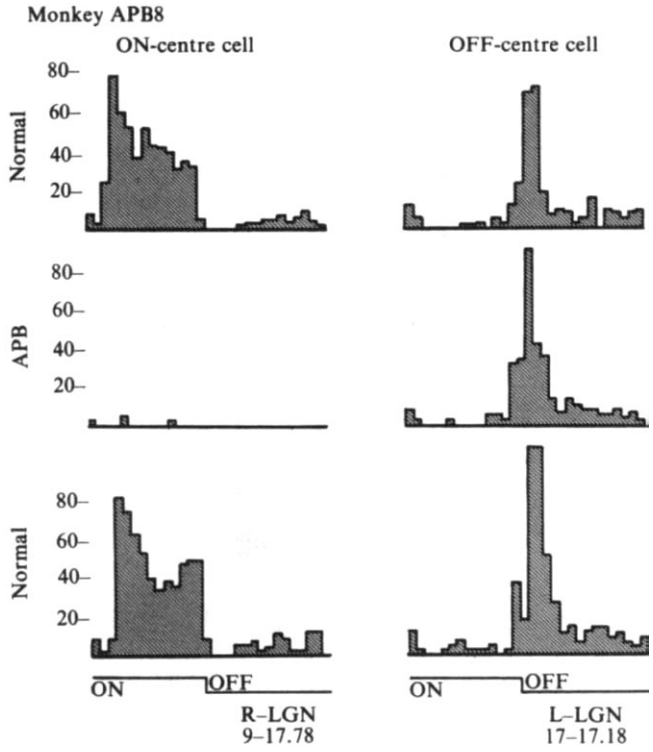


Fig. 1 The effect of retinal APB infusion into the left eye (200 μM APB for 1 min) on two simultaneously recorded single cells, one in layer 6 of the right LGN and the other in layer 2 of the left LGN. The receptive field centre of each cell was stimulated with a small spot of light, using two light beams, presented at the rate of 500 ms on and 500 ms off. Each histogram is based on responses elicited with 30 repeated stimulus presentations. The histograms on top were obtained before APB infusion, those in the middle after the infusion had blocked the retinal on responses, and those on the bottom after the ON-centre cell had recovered. APB abolished the responses of the ON-centre cell and had no significant effect on the OFF-centre cell.

allowed us selectively to infuse a solution which in addition to the perfusate contained 100–800 μM of APB. Electrodes were placed in the LGN and striate cortex of 10 animals. In most cases we recorded simultaneously with two electrodes, either with one in each LGN, or with one in the LGN and the other in the striate cortex. Stationary, flashing spots or moving light bars were presented on a tangent screen with a computer-driven optic bench[9]. The effects of APB infusion were studied at 124 LGN (78 single cells) and 55 cortex sites (41 single cells); their receptive fields were within 12° of the fovea. While recording in the LGN it was often also possible to monitor the activity of retinal ganglion cell axons[10].

In the LGN, brief 1–3-min infusion of 100–200 μM APB effectively blocked the activity of ON-centre cells within 1–3 min. In successful preparations, parvocellular ON-centre cells, most of which have colour-opponent receptive field properties[11–13], were blocked the whole time. Magnocellular ON-centre cells, which have broad-band receptive field properties[11–13], were more resistant to APB infusion; their responses were generally blocked several seconds later than those of parvocellular cells, and in some animals the response was reduced but not entirely eliminated. Therefore, to be certain about the extent to which ON responses were blocked while we studied ON–OFF interactions, we always assessed the responses of cells in the magnocellular lamina. The data reported here are based on those cases where we established that both the magnocellular and parvocellular responses were blocked by our infusions. For the most part, OFF-centre cells were not significantly affected by the APB. Figure 1 shows the effects of APB infusion on an ON-centre and an OFF-centre cell recorded simultaneously with two electrodes in the LGN. The infusion abolished the responses of the ON-centre cell and had little effect on the OFF-centre cell. Recordings from retinal ganglion cell axons showed similar effects.

To determine the extent to which the centre-surround antagonism observed in LGN cells is produced by the convergence of the ON and OFF channels, we selectively stimulated the centre and surround of LGN cells before and during APB infusion. Figure 2 shows one of the several methods we used. The receptive field centre was stimulated with a spot of light

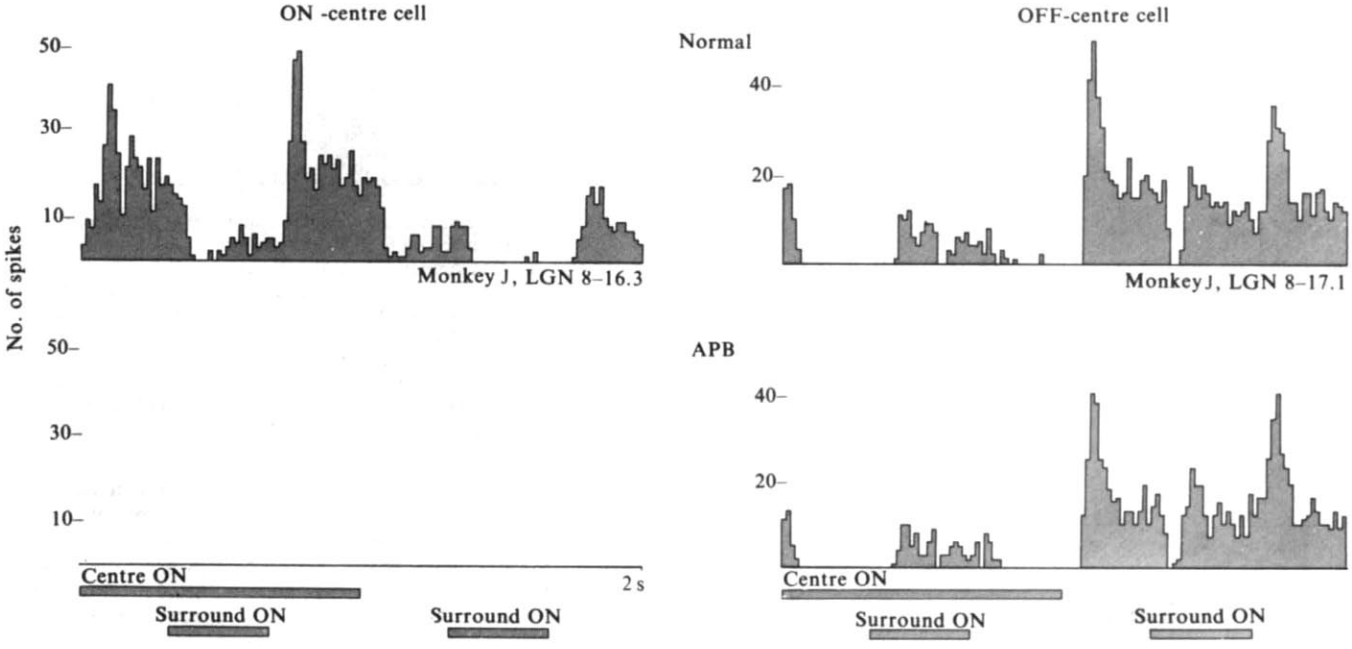


Fig. 2 The responses of an ON-centre and an OFF-centre colour-opponent cell to centre and surround stimulation before and after 200 μM infusion of APB into the eye. The centre stimulus (0.75° green spot for the ON-centre cell and 0.75° red spot for the OFF-centre cell) was flashed for 1 s every 2 s. The surround stimulus (9° red spot for the ON-centre cell and 9° green spot for the OFF-centre cell) was flashed on for 360 ms twice during each cycle, first 320 ms after the centre spot appeared and second 320 ms after the centre spot was extinguished. When the surround stimulus appeared alone, it yielded a response opposite in sign to that produced by the centre stimulus. Each histogram is based on 30 cycles. APB infusion blocked all the responses elicited from the ON-centre cell and had no significant effect on the responses of the OFF-centre cell.

© 1982 Nature Publishing Group

flashed on for 1 s every 2 s. The surround was stimulated with a large spot of a different wavelength, which appeared once with the centre spot on and once with the spot extinguished. Figure 2 shows the responses of an ON- and an OFF-centre cell for such conditions of stimulation. Both were colour-opponent cells, and were stimulated accordingly. The centre of the first cell was activated with a small green spot, while the surround was stimulated with a large red spot. The opposite arrangement was used for the OFF-centre cell, which for its centre required a red stimulus and for its surround a large green spot.

In the ON-centre cell, the surround stimulus, when paired with the centre spot, caused inhibition; when it was presented alone, it yielded a response opposite in sign to that produced by the centre stimulus. APB infusion (200 μM for 1 min) eliminated all the responses of the ON-centre cell, including the OFF response produced by surround stimulation. Similar infusion did not have a differential effect on the centre and surround responses of the OFF-centre cell; the ON response produced by the surround stimulation while the centre spot was on persisted. The centre-surround organization of OFF-centre broad-band cells was also unaffected by APB infusion. These findings suggest that the enhanced centre-surround antagonism of LGN cells, evident in both cat and monkey[11,12,14], is not produced, in the monkey at least, by interactions between the ON and OFF channels, as some investigators have suggested for the cat[14,15].

In the striate cortex we examined the effects of APB infusion on the light- and dark-edge responses of cortical cells, their directionality and their orientation specificity[9,16]. When the ON responses were blocked by APB infusion in both the parvocellular and magnocellular layers of the LGN, the light-edge response of cortical cells was eliminated. An example obtained from a complex cell is shown in Fig. 3. The cell was stimulated with a moving light bar, which elicited a response to both the leading light edge and the trailing dark edge. Infusion of 200 μM APB eliminated the light-edge response, suggesting that it is produced by the input from the ON channel. In some cortical cells more complicated interactions were evident. Figure 4 shows one such example. This simple cell[9], when stimulated in the same fashion as the cell shown in Fig. 3, responded only to a light edge. Infusion of 100 μM of APB eliminated the light-edge response, and, unexpectedly, uncovered a dark-edge response. Part of the way to recovery, both responses were

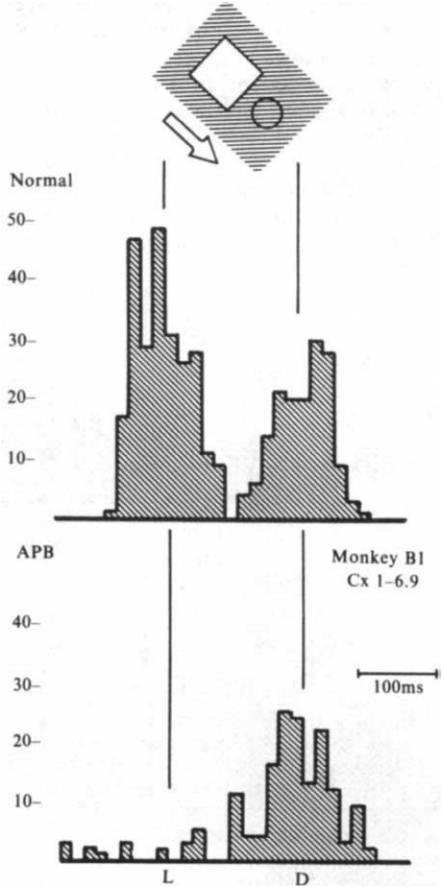


Fig. 3 The responses of a complex cell in striate cortex to a moving 1° wide light bar, as shown in the inset, before and after the infusion of 200 μM APB into the eye. The stimulus was moved across the receptive field at 4° per s. Data were collected for 30 repeated trials. The vertical lines show expected location of the light-edge (L) and dark-edge (D) responses. APB blocked the light-edge response and had little effect on the dark-edge response.

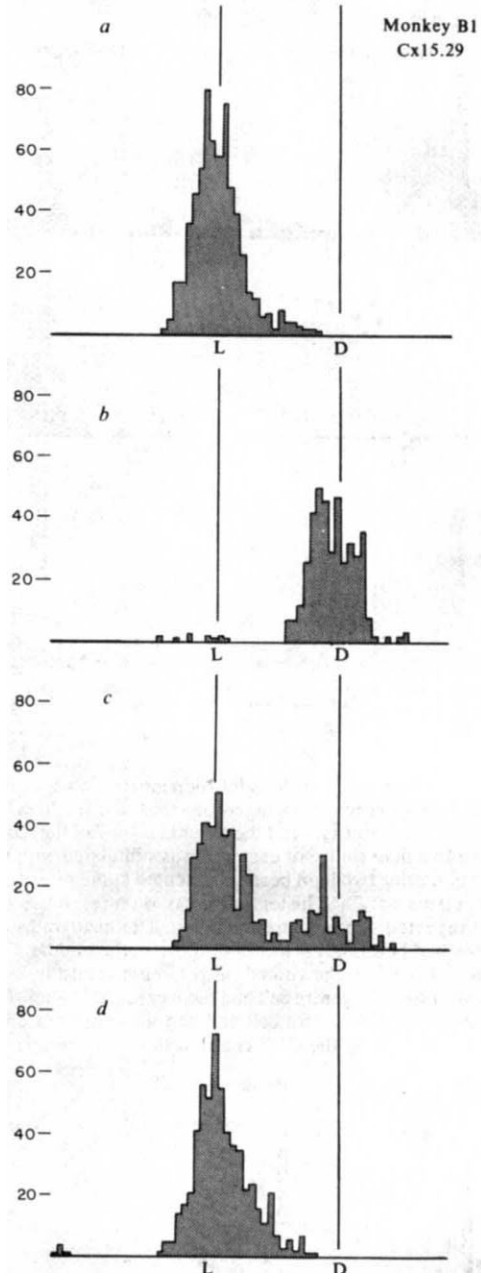


Fig. 4 The responses of a simple cell in the striate cortex before and after the infusion of APB. A 1° wide light bar was swept across the receptive field from left to right at 2° per s. Each histogram is the result of 30 repeated sweeps. The vertical lines show the expected location of the light-edge (L) and dark-edge (D) responses. *a*, Normal response during infusion of control perfusate; *b*, response after infusion of 100 μM APB had taken effect; *c*, partial recovery after termination of APB infusion; *d*, full recovery following termination of APB infusion. APB infusion uncovered a dark-edge response for this cell.

© 1982 Nature Publishing Group

evident. Recordings in magnocellular LGN have shown that APB infusion fully blocked the ON responses. These observations suggest that in addition to convergent excitation, significant inhibitory interactions can also occur between the ON and OFF channels in the striate cortex.

The orientation and direction specificities of cortical cells were assessed by examining the responses of cells to light and dark edges which were moved across receptive fields in various orientations in random order before, during and after blocking the retinal ON system. The directionality and orientation selectivity of the cortical cells we studied were unaltered by APB infusion, suggesting that these attributes are not produced by interactions between the ON and OFF channels[17]. Work carried out concurrently by J. C. Horton and H. Sherk using a different procedure has yielded similar results in the cat[18].

Thus, my results suggest that in the rhesus monkey ON-centre and OFF-centre retinal ganglion cells drive predominantly ON-centre and OFF-centre LGN cells respectively. The surround inhibition observed in LGN cells does not appear to be a product of the interaction between the ON and OFF channels. In the striate cortex the ON and OFF channels converge, with the light-edge response often produced by the input from ON-centre cells and the dark-edge response produced by OFF-centre cells. Inhibitory interactions between these channels are evident in the cortex. The orientation and direction specificity of cortical cells were unaffected by blocking the ON system, indicating that each channel has access to the mechanisms responsible for these attributes.

I thank Denice Couch-Helwig, Martha Rhodes and Nancy Stec for their assistance. This work was supported in part by NIH grants EY 02621 and EY 00676 and NSF grant BNS 8019714.

Received 13 October 1981; accepted 31 March 1982.

1. Slaughter, M. M. & Miller, R. F. *Science* **211**, 183 (1981).
2. Cleland B. G. & Levick, W. R. *J. Physiol., Lond.* **240**, 421 (1974).
3. DeMonasterio F. M. & Gouras, P. *J. Physiol., Lond.* **251**, 167 (1975).
4. Enroth-Cugell, C. & Robson, J. G. *J. Physiol., Lond.* **187**, 517 (1966).
5. Kuffler, S. W. *J. Neurophysiol.* **16**, 37 (1953).
6. Schiller, P. H. & Malpeli, J. G. *J. Neurophysiol.* **40**, 428 (1977).
7. Stone, J. & Fukuda, Y. *J. Neurophysiol.* **37**, 722 (1974).
8. Cunningham, R. & Miller, R. F. *Brain Res.* **117**, 341 (1976).
9. Schiller, P. H., Finlay, B. L. & Volman, S. F. *J. Neurophysiol.* **39**, 1288 (1976).
10. Schiller, P. H. & Malpeli, J. G. *Brain Res.* **137**, 387 (1977).
11. Dreher, B., Fukuda, Y. & Rodieck, R. W. *J. Physiol., Lond.* **258**, 433 (1976).
12. Schiller, P. H. & Malpeli, J. G. *J. Neurophysiol.* **41**, 788 (1978).
13. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **29**, 1115 (1966).
14. Hubel D. H. & Wiesel, T. N. *J. Physiol., Lond.* **155**, 383 (1961).
15. Maffei, L. & Fiorentini, A. *J. Neurophysiol.* **35**, 65 (1972).
16. Hubel, D. H. & Wiesel, T. N. *J. Physiol, Lond.* **195**, 215 (1968).
17. Hubel, D. H. *J. opt. Soc. Am.* **53**, 58 (1963).
18. Horton, J. C. *Neurosc. Abstr.* **7**, 11.8 (1981).

Electrophysiological action of kainic acid and folates in the *in vitro* olfactory cortex slice

C. R. Auker*, D. J. Braitman† & S. L. Rubinstein

Physiology Department, Armed Forces Radiobiology Research Institute, Bethesda, Maryland 20814, USA

Evidence has mounted that kainic acid (KA), a powerful neuroexcitatory and neurotoxic agent, acts at a specific class of receptors distinct from those mediating the excitatory actions of glutamate[1–3]. This has prompted a search for the endogenous ligand for the KA-specific receptors. Recently, Ruck *et al.*[4] reported that the naturally occurring folic acid derivative methyltetrahydrofolate (MTHF) is a potent and specific competitor for KA-binding sites in rat cerebellar membranes, having one-tenth the binding activity of KA at these sites. They suggested that MTHF may be the endogenous KA receptor ligand. Olney *et al.*[5] found that injections of various folates—MTHF, folic acid (pteroyl-L-glutamic acid, PGA), or folinic acid (formyltetrahydrofolate, FTHF)—into rat amygdala and striatum produced KA-like seizures and associated brain damage at sites distant from the site of injection. However, these folates did not mimic the direct neurotoxic effects of KA at the injection site. Furthermore, Roberts *et al.*[6] reported that MTHF injected into rat cerebellum produced slower and less complete neuronal degeneration than did KA and, unlike KA and glutamate, did not increase cerebellar cyclic GMP levels. In light of the contradictory results from these binding, histological and biochemical studies, we have tested the electrophysiological equivalency of KA and the folates in rat olfactory cortex slices. KA is a potent agonist at the receptors of the terminal synapses of the lateral olfactory tract (LOT); in fact, the natural transmitter at this synapse appears to act at a KA-preferring receptor[7]. Also, olfactory cortex is extremely sensitive to the neurotoxic action of KA[8]. We report here that MTHF, FTHF and PGA had little or no effect on LOT-stimulated field potentials even at millimolar concentrations, whereas KA exhibited excitatory effects at concentrations as low as 5×10^{-7} M. It therefore seems unlikely that MTHF or the other folates are endogenous ligands for the KA receptor in olfactory cortex.

* Present address: Ophthalmology Department, Georgetown University Medical Center, Reservoir Road, Washington, DC, USA.
† To whom correspondence should be addressed.

Thin slices (300–450 μm) of rat olfactory cortex were cut and incubated as previously described[7]. Each slice was transferred to a recording chamber that was constantly perfused with oxygenated normal Ringer's solution. The LOT was stimulated with double shocks of 20–50 V intensity and 50 μs duration via a tungsten bipolar electrode placed across the tract. Paired

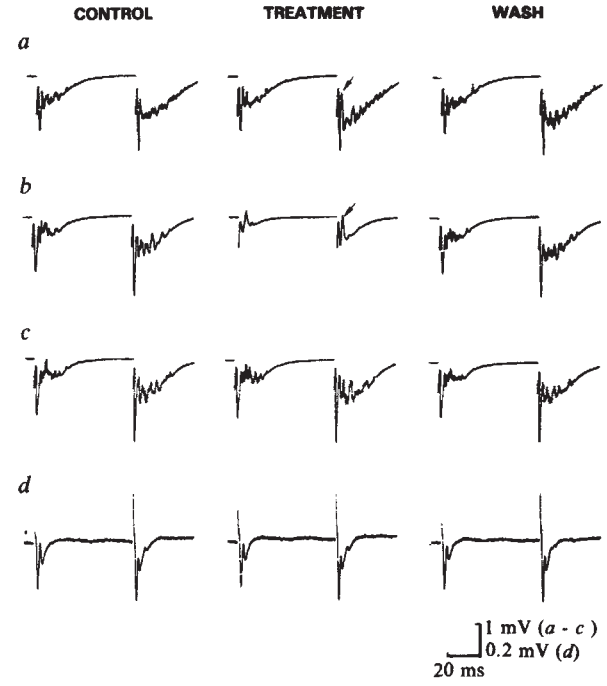


Fig. 1 Effects of KA and folates on LOT-stimulated field potentials in rat olfactory cortex slices. Each trace shows two field potentials in response to double shock stimulation of LOT. As previously reported, the second field potential of each pair is larger than the first[7,10]. Initial downward deflection in each trace is the beginning of the e.p.s.p.[11]. Rapid upward deflection is due to the superimposition of the population spike onto the slow wave component of the field potential[10,11]. The arrows in *a* and *b* indicate the population spike in the second field potential. Spiking in the late wave of the field potential indicates asynchronous neural firing. Slices were superfused with: *a*, 5×10^{-7} M KA for 15 min; *b*, 10^{-5} M KA for 5 min; *c*, 10^{-4} M FTHF for 15 min; and *d*, PGA for 20 min. The duration of wash was equal to or greater than duration of treatment. FTHF was prepared in oxygenated Ringer's just before use and was protected from light during experiment.

© 1982 Nature Publishing Group

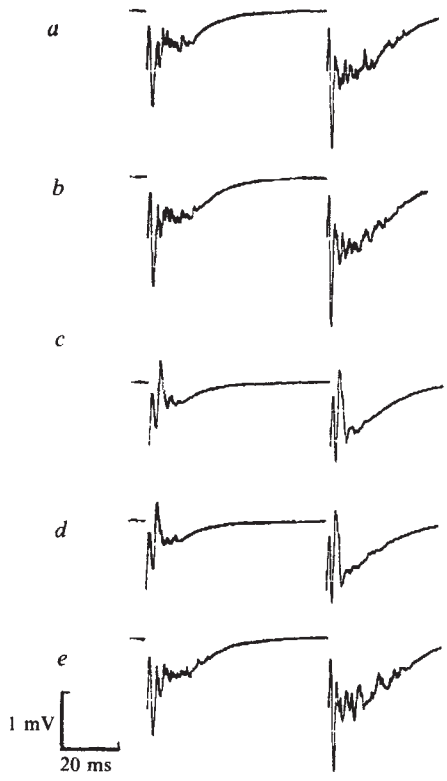


Fig. 2 Lack of agonist and antagonist activity of MTHF in rat olfactory cortex slice. *a*, Control: cortical field potentials evoked by double shock to LOT. Effect of treatment *b*, for 10 min with 10^{-4} M MTHF; *c*, for 8 min with 5×10^{-6} M KA. *d*, Effects of combined treatment for 7 min and 30 s with both 10^{-4} M MTHF and 5×10^{-6} M KA. *e*, Wash for 31 min. MTHF was prepared in oxygenated Ringer's just before use and was protected from light during experiment.

shocks were separated by 60 ms and administered every 4 s. Shock intensities were adjusted to produce field potentials of maximum amplitude. Glass micropipettes with ~50 μm cut tips were filled with normal Ringer's solution and used to record field potentials from the pial surface of the slice. After recording stable control field potentials for several minutes, the perfusate was changed to a Ringer's solution containing either KA, PGA, MTHF, FTHF, or a combination of KA and MTHF. Concentrations of these agents ranged from 10^{-8} to 10^{-3} M, and perfusion times from 5 to 30 min. This was followed by a wash in normal Ringer's for a period of time equal to or greater than that of the drug treatment.

KA administered in concentrations $>10^{-8}$ M exhibited dose-dependent excitatory effects on the LOT-stimulated field potential. The lowest effective bath concentration, 5×10^{-7} M KA, increased the amplitude of the population spike, indicating an increase in synchronous neuronal firing, but had little effect on the slow wave representing the population excitatory post-synaptic potential (e.p.s.p.) (Fig. 1*a*). At higher concentrations (Figs 1*b*, 2*c*), KA application resulted in an initial increase in the width and amplitude of the population spike followed by a decrease in amplitude of the e.p.s.p. As the e.p.s.p. approached the point of abolishment, the population spike increased to maximum size and then decreased and disappeared. These events appear to represent excitation followed by depolarization block. The actions of KA were reversible if wash with normal Ringer's was initiated no later than the time of disappearance of the population spike. In contrast to KA, PGA (Fig. 1*d*), FTHF (Fig. 1*c*) and MTHF (Fig. 2*b*) had little or no effect on the LOT-stimulated e.p.s.p. or population spike even when administered in concentrations as high as 1 mM. At concentrations $\geq10^{-4}$ M, these substances occasionally increased asynchronous spiking during the late stage of the field potential, an action consistent with the weak excitatory effects demonstrated for PGA and FTHF in cat cortex[9]. As none of the folates reproduced the effects of KA (even at 2,000 times the KA concentration), it seems unlikely that the folates are physiological ligands for the KA receptor.

As MTHF has been reported to be a potent competitor for KA-binding sites in rat cerebellum[6] but does not seem to be an agonist in rat olfactory cortex, we tested whether MTHF is an antagonist at the KA receptors. MTHF (10^{-4} M) and KA (5×10^{-6} M) were individually and then simultaneously bath-applied to the slice during LOT stimulation. KA alone decreased the amplitude of the e.p.s.p. and increased the amplitude of the population spike, whereas MTHF alone had no effect; in combination with KA, MTHF neither diminished nor slowed the effects of KA on the e.p.s.p. or the population spike (Fig. 2). Thus MTHF does not appear to be an antagonist at the KA receptor in olfactory cortex.

Evidence exists that the folates can excite central nervous system neurones. Davies and Watkins[9] showed that cat cortical neurones were weakly excited by PGA and FTHF, and Loots *et al.* (cited in ref. 4) have shown that MTHF depolarizes frog spinal cord neurones. However, neither of these studies examined the relationship between the actions of the folates and KA. We have directly compared the electrophysiological effects of the folates and KA in a system that has been demonstrated to have KA receptors, and in which the endogenous transmitter has KA-like activity[7]. In this system, neither the folate derivatives (FTHF and MTHF) nor folic acid itself reproduced the effects of KA. Thus the folates cannot be considered to be the natural KA ligand at the terminal synapses of the LOT. The apparent discrepancy between our electrophysiological data and the binding data of Ruck *et al.*[4] may reflect (1) existence of different KA receptors in different brain regions, or (2) coexistence, in the same brain region, of at least two types of KA receptor, one that binds KA but not the folates and has electrophysiological effects, and another that binds KA and the folates but has no electrophysiological effects measurable by our techniques.

This work was supported by the Armed Forces Radiobiology Research Institute, Defense Nuclear Agency, under research work unit MJ 00005. The views presented in this paper are those of the authors. No endorsement by the Defense Nuclear Agency has been given or should be inferred.

Received 18 January; accepted 16 March 1982.

1. Hall, J. G., Hicks, T. D. & McLennan, H. *Neurosci. Lett.* **8**, 171–175 (1978).
2. Baudry, M. & Lynch, G. *Eur. J. Pharmac.* **57**, 283–285 (1979).
3. Watkins, J. C. *Adv. Biochem. Psychopharmac.* **29**, 205–212 (1981).
4. Ruck, A., Kramer, S., Metz, J. & Brennan, M. J. W. *Nature* **287**, 852–853 (1980).
5. Olney, J. W., Fuller, T. A. & de Gubareff, T. *Nature* **292**, 165–166 (1981).
6. Roberts, P. J., Foster, G. A. & Thomas, E. M. *Nature* **293**, 654–655 (1981).
7. Hori, N., Auker, C. R., Braitman, D. J. & Carpenter, D. O. *Cell. molec. Neurobiol.* **1**, 115–120 (1981).
8. Olney, J. W. & de Gubareff, T. in *Kainic Acid as a Tool in Neurobiology* (eds McGeer, E., Olney, J. W. & McGeer, P. L.) 201–217 (Raven, New York, 1978).
9. Davies, J. & Watkins, J. C. *Biochem. Pharmac.* **22**, 1667–1668 (1973).
10. Richards, C. D. *J. Physiol., Lond.* **222**, 209–231 (1972).
11. Richards, C. D. & Sercombe, R. *J. Physiol., Lond.* **211**, 571–584 (1970).

Neurone differentiation in cell lineage mutants of *Caenorhabditis elegans*

J. G. White*, H. R. Horvitz† & J. E. Sulston*

* MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK
† Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

The nematode *Caenorhabditis elegans* develops by an essentially invariant sequence of cell divisions[1–3] leading to an adult complement of 959 somatic cells. In this organism cell fate is correlated with cell lineage, suggesting that genealogy may be a determining factor for the differentiated state of a cell. The study of mutants with altered cell lineages may help elucidate the precise mechanisms by which cell fate is decided. Several

0028–0836/82/240584—04$01.00
© 1982 Macmillan Journals Ltd

cell lineage mutants have been isolated and characterized[4,5], some having more and some fewer cell divisions than wild type. We have now investigated the cell types produced by two cell lineage mutants; these mutants exhibit blocks in certain terminal or near terminal cell divisions, which in normal animals generally give rise to daughter cells that differentiate into distinctly different cell types. We find that the blocked cells in the mutants generally exhibit the differentiated characteristics of only one of the two daughter cells that normally would be produced. The differentiated state of the blocked precursors may be due to an intrinsic dominance of one cell type over another in what is essentially a fused cell, and/or it may reveal the state of commitment of the precursor in wild-type animals.

We have studied two very similar mutants *unc*-59(*e*1005) and *unc*-85(*e*1414), both of which are variably blocked in some of the later divisions of the lineages that produce the adult complement of motoneurones in the ventral nerve cord[4,5]. These lineages have been characterized in wild-type animals by following the development of living animals in the light microscope[1]. The nuclei of 12 precursor cells (designated P1–P12) migrate into the ventral cord during the first larval stage; each cell then divides to produce a neuroblast and a hypodermal cell. The 12 neuroblasts (together with an additional neuroblast present at hatching) undergo identical sequences of divisions (Fig. 1), each producing 5 cells that intercalate with the pre-existing juvenile motoneurones to form a single row of cells along the ventral cord. All cleavages have longitudinal spindle axes, and descendant cells maintain their relative antero-posterior positions throughout development. Certain cells derived from some of the neuroblasts die soon after their formation; the pattern of cell death is invariant.

The structure of the ventral cord of wild-type animals has been deduced by reconstructions from electron micrographs of serial sections[6]. In neuroblast lineages that do not contain cell deaths, the five cells produced differentiate into five distinct classes of motoneurone designated VA, VB, VC, AS and VD (Fig. 1). Classes are defined on the basis of patterns of synaptic connections (see Pn, Table 2) and morphology (Fig. 2). Either criterion is sufficient to assign a cell to a class.

The post-embryonic development of ventral cord cells was observed in several *unc-59* and *unc-85* animals. In these mutants, P-cell-derived neuroblasts produced fewer progeny cells than normal (Fig. 1), because some cell divisions failed. The set of affected cells varied between animals. Failures resulted in polyploid cells; sometimes the nuclei failed to divide after DNA replication, and sometimes apparently normal nuclei formed but then fused to form a tetraploid nucleus or remained separated in a binucleate cell[5]. *unc-59* and *unc-85* individuals that displayed several examples of undivided ventral cord nuclei were selected for reconstruction from electron micrographs of serial sections.

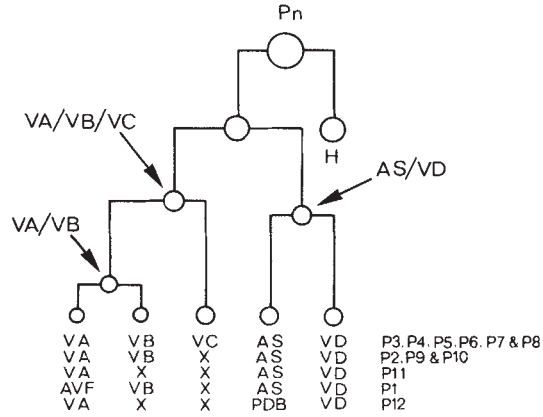


Fig. 1 The ventral cord precursor cells P1–P12 are hypodermal cells (J.G.W., unpublished observations). Their nuclei migrate into the ventral cord and each divides to produce a ventral hypodermal cell (H) (which functionally replaces the mother) and a neuroblast. All 12 neuroblasts then undergo identical series of divisions to produce five descendants[1]. All spindle axes are longitudinal; in the diagram anterior daughters are drawn to the left and posterior daughters to the right. The fates of the descendants from each precursor are shown. Cell types have been assigned on the basis of cell morphology and synaptic connectivity[6]. Precursors P3–P8 produce five classes of ventral cord motoneurone (VA, VB, VC, AS, VD). In the P1–P2 and P9–P12 lineages, cells lineally equivalent to VC neurones undergo programmed cell death (X), as do cells lineally equivalent to VB neurones derived from P11–P12. Alternative fates are seen for some of the cells derived from the two precursors at the ends of the cord (P1, P12): P1 produces an AVF interneurone instead of a VA, and P12 produces a PDB interneurone instead of an AS. (AVF and PDB are two distinct classes of interneurone which are quite dissimilar to the motoneurone classes.) These alternative fates are probably specified by local interactions[1,14]. In the *unc-59* and *unc-85* animals, certain cells which divide in the wild-type fail to do so. Such cells are labelled according to the fate of their normal descendants; for example, an AS/VD cell should have divided to produce an AS and a VD neurone.

Table 1 Post-embryonic development of P cells from *unc-85* and *unc-59* animals

Mutant	Blocked precursor		Cell type
unc-85	P3	VA/VB/*	VB
	P3	AS/VD	VD
	P4	VA/VB/*	VB
	P5	VA/VB	VB
	P6	VA/VB/*	VB‡
	P6	AS/VD	VD
unc-59			
Animal 1	P3	VA/VB/VC	VB
	P3	AS/VD	VD
Animal 2	P8	VA/VB	VA
	P8	AS/VD	VD
	P9	VA/VB/†	VA
	P9	AS/VD	VD
	P10	VA/VB/†	VA
	P10	AS/VD	VD
	P11	VA/†/†	VA‡
	P11	AS/VD	VD

*In these lineages a VC nucleus was produced but was not present when the adult was reconstructed. These nuclei probably either fused to make a VA/VB/VC or died.

† The cells lineally equivalent to VC in the P1–P2 and P9–P12 lineages and to VB in the P11–P12 lineages normally undergo programmed cell death[1].

‡ Axons failed to grow out of these cells although they had the synaptic contacts characteristic of the cell types indicated.

In both *unc-59* and *unc-85* animals all blocked cells were found to have differentiated into neurones. Most of these neurones could be unequivocally identified as belonging to one of the motoneurone classes normally produced from these lineages; specifically, the blocked cells acquired the differentiated characteristics of one of the daughter cells that would normally be produced. These results are summarized in Table 1. All blocked VD/AS precursors differentiated into cells that had all the distinguishing characteristics of VD motoneurones (Fig. 2, Tables 1, 2). VA/VB or VA/VB/VC precursors differentiated into VB-like cells in the anterior ventral cord and into VA-like cells in the posterior ventral cord (Tables 1, 2). The choice of VA versus VB for the blocked precursors may be related to the regional preference exhibited by the progeny of these precursors in wild-type development: a VB but no VA is produced by the anterior P1 lineage, whereas a VA but no VB is produced by the posterior P11 and P12 lineages (Fig. 1).

Several differences were apparent between the blocked precursors and the corresponding cell types in wild-type animals: the blocked cells had more synaptic inputs (Table 1), larger nuclei and extra branches (Fig. 2). These extra branches were generally devoid of synapses and had none of the characteristics of the other cell type. Even more extensive branching has been seen from precursors that are blocked earlier in ventral cord lineages in the mutant *lin-5* (ref. 7). The supernumerary branches may be a consequence of the polyploid nature of these cells; perhaps the total process length produced from a given neurone is proportional to its ploidy. The embryonically produced neurones in this region (that is, the DA, DB and DD motoneurones[6,8]) were normal in morphology and connectivity.

The observation that characteristics of some cell types were expressed normally in the blocked cells implies that the normal

© 1982 Nature Publishing Group

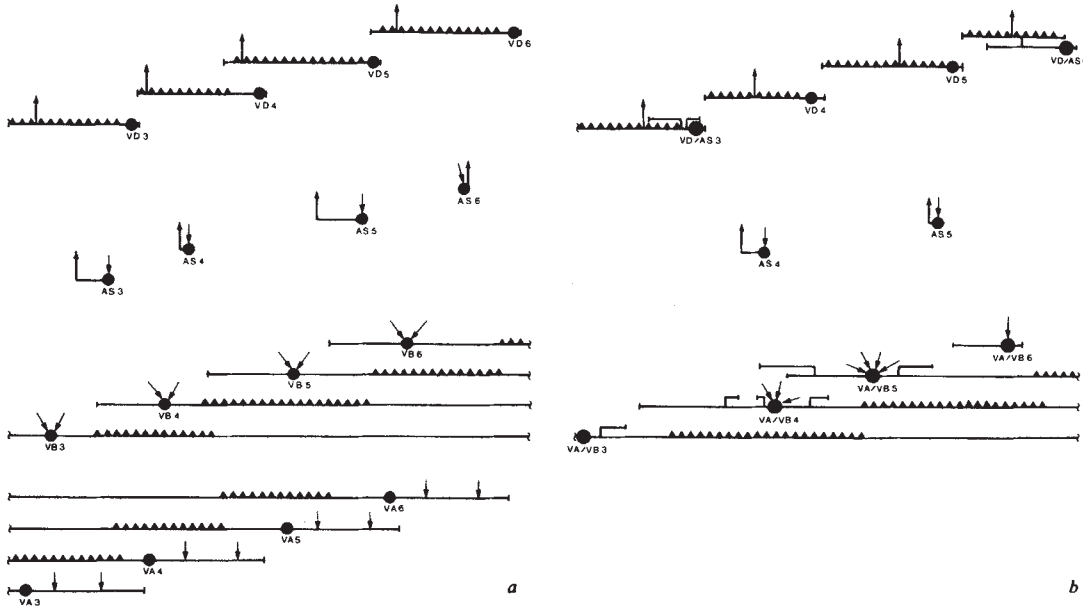


Fig. 2 The morphologies of the motoneurones derived from the precursors P3–P6 are shown for wild type (*a*) and *unc-85* (*b*). The wild-type structures are as described in ref. 6, and the mutant structures were derived, using similar techniques, by reconstruction from 3,000 serial section electron micrographs from an animal of known lineage. Four of the five motoneurone classes derived from the P cells are shown: VA, VB, AS and VD (class VC has been omitted, because it has few distinguishing characteristics). The extent of the processes in the ventral cord and the positions of the cell bodies (–●–), neuromuscular junctions (NMJs) (▲▲▲), commissures (|→) and synaptic input (|←) are shown. The horizontal axis is equivalent to the longitudinal axis of the animal (anterior is drawn to the left); the vertical axis has been used simply to separate the neurones into their respective classes. VA neurones, which have anteriorly directed axons, form a regular sequence of regions of motor activity with little or no overlap between adjacent members of a class. VB neurones are similar but have posteriorly directed axons. VA neurones receive their synaptic input in a dendritic region near the cell body on a branch opposite the axon; VB neurones receive their synaptic input predominantly at the cell body. The axons from class AS motoneurones lead to the dorsal cord via commissures and innervate dorsal muscles. They have a short dendritic region in the ventral cord. The axons from VD neurones end abruptly in gap junctions to axons from adjacent VD neurones. The dendritic regions behave in a similar fashion in the dorsal cord (not shown) and connect to the ventral cord via commissures. In the *unc-85* animal, the VA/VB-blocked precursors from P3–P5 had morphologies similar to wild-type VB neurones, forming NMJs from posteriorly directed axons. The blocked VA/VB precursor from P6 had failed to grow an axon. The AS/VD precursors from P4 and P5 had divided normally, and each produced apparently normal AS and VD daughters. The blocked AS/VD cells from P3 and P6 had the characteristic morphologies of VD motoneurones with their processes forming NMJs and ending abruptly in gap junctions with adjacent normal VD motoneurones. The only differences between the morphologies of these blocked precursors and those of their normal posterior daughters is that the former have larger nuclei, a few extra branches and more synaptic inputs.

complement of wild-type cell divisions is not a necessary prerequisite for the differentiation of these cell types; however, some other cell cycle event (such as DNA replication[15]) may be required. Blast cells that are blocked earlier in a lineage either by drugs[9,10] or in the mutant *lin-5* (ref. 7), also exhibit characteristics of their normal descendants, but in these cases characteristics of several cell types may be present in one polyploid cell. This may be a result of either localized differentiation or excessive dilution of regulatory elements within these abnormally large, highly polyploid cells. The isolation and characterization of mutants that are blocked at intermediate levels may shed some light on these differences.

Not all late divisions failed in these mutants, and those that did not fail produced normal cells. This observation indicates that the suppression of the differentiated characteristics of certain cell types seen in the blocked precursors does not simply

Table 2 Number of synapses formed by blocked precursors and typical synapses of wild-type ventral cord neurones (Pn)[6]

		AVA	AVB	VA	VB	VC	AS	VD	NMJ	
unc-85										
P3	VA/VB/*	‡	‡						15	
P3	AS/VD				§ 1			= 1	19	
P4	VA/VB/*		= 6						20	
P5	VA/VB	= 1	= 9		= 5				8	
P6	VA/VB/*		= 1		= 5					No axon
P6	AS/VD				§ 3			= 1	14	
unc-59 (1)										
P3	VA/VB/VC		= 15					# 1	5	
P3	AS/VD							= 6	14	
unc-59 (2)										
P8	VA/VB	§ 6 = 6							‡	
P8	AS/VD			§ 3				= 2	11	
P9	VA/VB/†	§ 3 = 5						# 3	22	
P9	AS/VD			§ 3				= 2	13	
P10	VA/VB/†	§ 9 = 5						# 3	16	
P10	AS/VD							= 2	10	
P11	VA/†/†	§ 7								No axon
P11	AS/VD							= 2	12	
Wild type										
Pn	VA	§ 4 = 3					= 1	# 1	16	
Pn	VB		= 2		= 2			# 2	9	
Pn	VC					§ 3 = 5		# 7	2	
Pn	AS	§ 2 = 1	§ 2	= 1						
Pn	VD			§ 1	§ 2	§ 7		= 1	26	

Values for wild type are the mean values from precursors P3 to P5 obtained from a reconstruction of a wild-type animal. #, Presynaptic; §, postsynaptic; =, gap junctions.

* In these lineages a VC nucleus was produced but was not present when the adult was reconstructed. These nuclei probably either fused to make a VA/VB/VC or died.

† The cells lineally equivalent to VC in the P1–P2 and P9–P12 lineages and to VB in the P11–P12 lineages normally undergo programmed cell death[1].

‡ Processes in regions that normally would display these synapses were out of the region of reconstruction.

© 1982 Nature Publishing Group

reflect the inability of the mutants to produce these cell classes but implicates the failure to divide as the cause of their suppression.

The lack of expression of one cell type in the undivided, probably tetraploid precursors of *unc-59* and *unc-85* may be because there is an intrinsic mutual exclusivity in the expression of cell type such as has been described for certain fused cells[11] and cultured sympathetic neurones[12]. An alternative interpretation of these observations is that the differentiated state of the precursors reflects the state of commitment of these cells in wild-type animals; thus, one of the daughters would normally inherit the state of commitment of the mother. Some support for this notion is provided by observations of mutants in the genes *unc-86* and *lin-4*; these mutants undergo extra cell divisions in specific lineages. In these cases one of the daughters reiterates the characteristic divisions of the mother in a stem cell manner[13], indicating that it has the same state of commitment as its mother cell. These two interpretations are not necessarily incompatible if cells may only express one state of commitment at any time because of an intrinsic mutual exclusivity of cell states.

The electron microscopy was done by N. Thomson and Marilyn Anness. We thank S. Brenner and W. Fixsen for helpful discussions. H.R.H. was supported by postdoctoral fellowships from the Muscular Dystrophy Associations of America and the USPHS and by USPHS grants GM24663 and GM24943.

Received 25 January; accepted 28 April 1982.

1. Sulston, J. E. & Horvitz, H. R. *Devl Biol.* **56**, 110–156 (1977).
2. Kimble, J. & Hirsh, D. *Devl Biol.* **70**, 396–417 (1979).
3. Deppe, U. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 376–380 (1978).
4. Horvitz, R. & Sulston, J. *Genetics* **96**, 435–454 (1980).
5. Sulston, J. & Horvitz, R. *Devl Biol.* **82**, 41–55 (1981).
6. White, J., Southgate, E., Thomson, N. & Brenner, S. *Phil. Trans. R. Soc.* B**275**, 327–348 (1976).
7. Albertson, D., Sulston, J. & White, J. *Devl. Biol.* **63**, 165–178.
8. White, J., Albertson, D. & Anness, M. *Nature* **271**, 746–766 (1978).
9. Whittaker, J. R. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2096–2100 (1973).
10. Laufer, J., Bazzicalupo, P. & Wood, W. *Cell* **19**, 569–577 (1980).
11. Davidson, R. L. in *Somatic Cell Hybridization* (eds Davidson, R. L. & de la Cruz, F.) 131–150 (Raven, New York, 1974).
12. Reichardt, L. & Patterson, P. *Nature* **270**, 147–151 (1977).
13. Chalfie, M., Horvitz, R. & Sulston, J. *Cell* **24**, 59–69 (1981).
14. Sulston, J. E. & White, J. G. *Devl Biol.* **78**, 577–598 (1980).
15. Satoh, N. & Susumu, I. *J. Embryol. exp. Morph.* **61**, 1–13 (1981).

Ovarian secretion of oxytocin is stimulated by prostaglandin

A. P. F. Flint & E. L. Sheldrick

ARC Institute of Animal Physiology, Babraham, Cambridge CB2 4AT, UK

Oxytocin induces regression of corpora lutea when administered to several domestic species[1–3] and a role for endogenous oxytocin in controlling the ovarian cycle is supported by the delay in luteal regression observed in sheep which have been immunized against oxytocin[4]. Although the posterior pituitary is the major source of oxytocin released at parturition and before suckling, recent evidence suggests that during the oestrous cycle the ovary may also secrete oxytocin. Circulating oxytocin and progesterone levels change in parallel during the oestrous cycle and after ovariectomy[5–7], and the corpus luteum[8] (but not other ovarian tissues [A.P.F.F. and E.L.S., unpublished]) contains high concentrations of oxytocin. We now extend this evidence by showing that in sheep, the concentration of oxytocin in ovarian venous blood exceeds that in arterial blood, and we find that ovarian secretion of oxytocin is stimulated by a prostaglandin $F_{2\alpha}$ analogue.

Experiments with Estrumate (cloprostenol, ICI 80996) have shown that when this potent prostaglandin $F_{2\alpha}$ analogue is used to cause luteal regression, the resulting decline in plasma progesterone concentration is accompanied by a drop in jugular venous concentrations of oxytocin during the 1–2 days following treatment (A.P.F.F. and E.L.S., unpublished). However, within 10 min of administering Estrumate there is a transient rise in plasma oxytocin which lasts up to 40 min. The finding that this transient increase is absent in ovariectomized sheep prompted us to measure the veno–arterial concentration difference for oxytocin across the ovary.

Anaesthesia was induced in four Clun Forest ewes 11–14 days after oestrus with intravenous pentobarbitone sodium and maintained after intubation with fluothane in oxygen. Mean oestrous cycle length in the flock was 16.5 days. Ovarian venous blood was collected from ovaries containing corpora lutea by inserting a polyvinyl catheter into an ovarian vein within 2 cm of the ovary and passing it towards the ovary. In two animals catheters were also inserted in ipsilateral utero-ovarian veins and directed away from the ovary. To allow collection of arterial blood and blood draining the posterior pituitary catheters were also inserted into a branch of the contralateral uterine artery and into the external jugular vein. Plasma was prepared from heparinized blood samples and oxytocin and progesterone levels measured by specific radioimmunoassays[6]. A high concentration of progesterone in the ovarian and utero-ovarian veins confirmed that blood was collected from an ovary bearing a corpus luteum.

Estrumate (125 μg, intramuscularly) caused a rapid (within 5 min) release of oxytocin from the ovary, which was maximal 10–20 min after treatment (Fig. 1). At peak oxytocin titres a veno–arterial concentration difference of up to 20 ng ml^{-1} was measured across the ovary (in an animal bearing two corpora

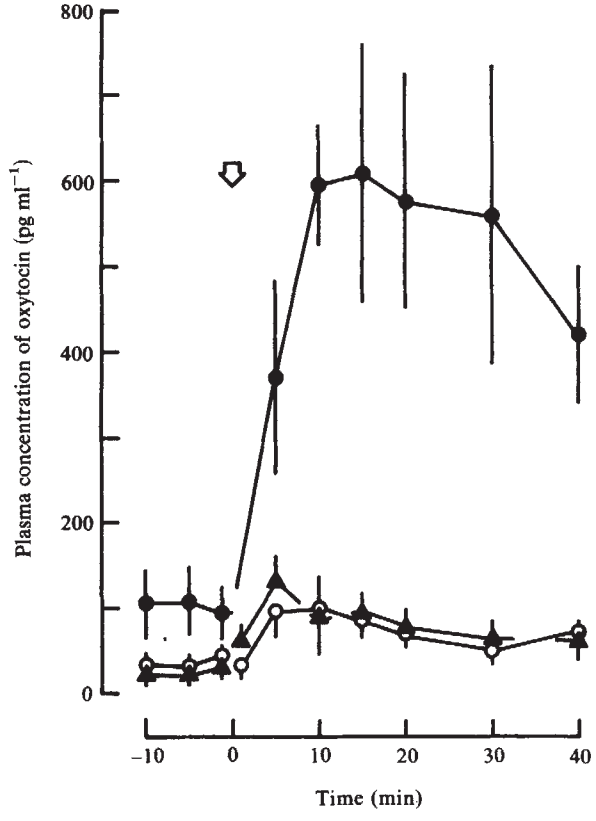


Fig. 1 Oxytocin concentrations measured by radioimmunoassay[6] in plasma from the ovarian vein (●), uterine artery (○) and jugular vein (▲) of four sheep 11–14 days after oestrus. Values are mean ±s.e.m. Ovarian venous concentrations from one animal bearing two corpora lutea in the sampled ovary have been halved; thus values are per corpus luteum. Estrumate (125 μg) was administered intramuscularly at $t = 0$ (arrowed). Anaesthesia was induced with pentobarbitone sodium and maintained with fluothane in oxygen. The antibody used in the radioimmunoassay showed negligible (<0.01%) cross-reactivity with arginine-vasopressin, vasotocin, luteinizing hormone-releasing hormone, ovine luteinizing hormone, progesterone and cloprostenol.

0028–0836/82/240587—02$01.00

© 1982 Macmillan Journals Ltd

lutea in the sampled ovary; data not included in Fig. 1); ovarian venous concentrations exceeded those in the utero-ovarian vein by a factor of 4.3 (mean of six samples taken simultaneously; $P < 0.005$ by analysis of variance). There was no detectable secretion by the pituitary either before or after administration of Estrumate; the mean concentration ratio jugular vein/artery for oxytocin was 1.1 ($n = 29$ samples). The concentration of oxytocin in blood draining the ovary exceeded the arterial concentration in samples taken before administering Estrumate by a factor of 2.8 ($n = 12$; $P < 0.001$), indicating basal ovarian secretion of oxytocin.

These observations show that oxytocin is released into the ovarian vein both before and after prostaglandin treatment; thus oxytocin is secreted by the ovary. It is not certain, however, whether the oxytocin released is synthesized in the ovary or sequestered in it from the blood stream (for example, bound to a receptor on the plasma membrane). The demonstration that oxytocin is released from the ovary before Estrumate administration must be interpreted with caution, as manipulation of the uterus during insertion of catheters may have caused sufficient uterine secretion of prostaglandin $F_{2\alpha}$ to stimulate ovarian release of oxytocin[9].

Assuming a minimum ovarian venous blood flow of 10 ml min^{-1}, and correcting ovarian venous concentrations for those in arterial blood, it can be calculated that during the 40 min following prostaglandin treatment average ovarian release of oxytocin was 0.24 μg per corpus luteum. This is approximately 20% of the luteal content of oxytocin reported by Wathes and Swann[8], and therefore it appears that a large proportion of the oxytocin contained in the ovary is available for release in response to prostaglandin stimulation. Although in the present experiments we used an analogue of prostaglandin $F_{2\alpha}$ to stimulate ovarian oxytocin secretion, it seems likely that oxytocin is also secreted in response to endogenous prostaglandin $F_{2\alpha}$ released from the uterus during the oestrous cycle. Synchronous surges of oxytocin–neurophysin, or of oxytocin itself, and of 13,14-dihydro-15-ketoprostaglandin $F_{2\alpha}$ (pulmonary metabolite of prostaglandin $F_{2\alpha}$) have been observed in sheep (ref. 10 and A.P.F.F. and E.L.S., unpublished) and oxytocin secretion has been reported in other species in response to native prostaglandin $F_{2\alpha}$[11,12].

Ovarian oxytocin may be involved in limiting luteal progesterone secretion by a local mechanism: oxytocin has been shown to inhibit steroidogenesis in ovarian[13] and testicular[14] tissues. It may also act systemically to augment the uterine secretion of prostaglandin $F_{2\alpha}$[15], thereby ensuring rapid completion of luteal involution, or lead to the reduction in luteal blood flow which occurs at luteolysis[16]: oxytocin has been reported to show vasoconstrictor properties in the mammary gland in some conditions[17]. However, as in so many cases of well established hormones being discovered in novel places[18], the physiological role of ovarian oxytocin remains to be determined.

Received 16 March; accepted 13 April 1982.

1. Armstrong, D. T. & Hansel, W. *J. Dairy Sci.* **42**, 533–542 (1959).
2. Milne, J. A. *Aust. vet. J.* **39**, 51–52 (1963).
3. Cooke, R. G. & Knifton, A. *Theriogenology* **16**, 95–97 (1981).
4. Sheldrick, E. L., Mitchell, M. D. & Flint, A. P. F. *J. Reprod. Fert.* **59**, 37–42 (1980).
5. Webb, R., Mitchell, M. D., Falconer, J. & Robinson, J. S. *Prostaglandins* **22**, 443–453 (1981).
6. Sheldrick, E. L. & Flint, A. P. F. *Prostaglandins* **22**, 631–636 (1981).
7. Schams, D., Lahlou-Kassi, A. & Glatzel, P. *J. Endocr.* **92**, 9–13 (1982).
8. Wathes, D. C. & Swann, R. W. *Nature* **297**, 225–227 (1982).
9. Roberts, J. S., Barcikowski, B., Wilson, L., Skarnes, R. C. & McCracken, J. A. *J. Steroid Biochem.* **6**, 1091–1097 (1975).
10. Fairclough, R. J. *et al. Prostaglandins* **20**, 199–208 (1980).
11. Ellendorff, F. *et al. Br. J. Pharmac.* **62**, 412P (1978).
12. Gillespie, A., Brummer, H. C. & Chard, T. *Br. med. J.* **1**, 543–544 (1972).
13. Tan, G. J. S., Tweedale, R. & Biggs, J. S. G. *J. Steroid Biochem.* **14**, xiii (1981).
14. Adashi, E. Y. & Hsueh, A. J. W. *Nature* **293**, 650–652 (1981).
15. McCracken, J. A. *Adv. Prostaglandin Thromboxane Res.* **8**, 1329–1344 (1980).
16. Pharriss, B. B., Cornette, J. C. & Gutknecht, G. D. *J. Reprod. Fert. Suppl.* **10**, 97–103 (1970).
17. Linzell, J. L. in *Lactation* Vol. 1 (eds Larson, B. L., & Smith, V. R.) 143–225 (Academic, New York, 1974).
18. Sharpe, R. M. *J. Reprod. Fert.* **64**, 517–527 (1982).

Racial differences in hypertension-associated red cell sodium permeability

Nina L. Etkin*, John R. Mahoney†, Michael W. Forsthoefel‡, James R. Eckman‡, John D. McSwigan†, Richard F. Gillum§ & John W. Eaton*†§

Departments of *Anthropology, †Laboratory Medicine and Pathology, and §Medicine, University of Minnesota, Minneapolis, Minnesota 55455, USA
‡ Department of Medicine, Emory University, Atlanta, Georgia 30322, USA

The prevalence and severity of essential hypertension varies greatly among human populations. Some of this variation is undoubtedly due to differences in dietary Na^+ levels[1]. However, even within groups which consume large amounts of salt, not all members develop increased blood pressure. Furthermore, the clinical expression of hypertension varies among races, and American blacks have particularly severe symptoms. These observations, in addition to a familial pattern in the occurrence of essential hypertension, suggest important genetic influences. Recent studies, primarily of northern European populations, have demonstrated abnormal erythrocyte Na^+ transport in affected individuals[2–5] and in approximately half the members of kindreds founded by a hypertensive progenitor[4,5]. We have found that red cell Na^+ transport can be assessed simply by measuring the unidirectional passive influx of $^{22}Na^+$ into ouabain-treated erythrocytes[5] and have applied this technique to American blacks with essential hypertension. The results suggest a complete absence of the abnormal erythrocyte Na^+ transport which is characteristic of white patients with essential hypertension. Thus, among American blacks, essential hypertension may have either a different genetic basis or a substantially different expression.

The white and black patients investigated had had supine diastolic blood pressures exceeding 95 mm Hg for at least 2 yr, and secondary hypertension was excluded by normal renal function and (in most cases) normal intravenous pyelogram. Normotensive controls had supine diastolic pressures of <85 mm Hg and negative family histories for essential hypertension. All patients were haemoglobin AA, as determined by isoelectric focusing.

Figure 1 shows the values obtained when erythrocyte Na^+ influx was measured in the four groups studied. In 21 patients with essential hypertension, erythrocyte $^{22}Na^+$ influx averaged twice that into red cells from 21 normotensive white controls (0.527 ± 0.128 compared with 0.266 ± 0.044 mmol Na^+ per l red cells per h respectively; $t = 8.87$, $P < 0.001$). The values for only 2 of 21 white patients with essential hypertension overlap the upper range of normal. This increased Na^+ influx is not influenced by antihypertensive medications administered to the patient[5] and is not caused by elevated blood pressure *per se*; erythrocytes from 12 patients with secondary hypertension (caused by renal disease) had values within the normal range (0.233 ± 0.060 mmol Na^+ per l red cells per h).

In contrast, red cells from 32 American black patients with essential hypertension showed no evidence of enhanced Na^+ influx. In these patients, values for erythrocyte Na^+ influx (0.202 ± 0.059 mmol Na^+ per l red cells per h) were indistinguishable from those for 23 normotensive black controls (0.216 ± 0.069 mmol Na^+ per l red cells per h; $t = 0.79$, $P > 0.10$). There is, however, a highly significant difference between hypertensive whites and blacks in erythrocyte Na^+ influx ($t = 10.66$, $P < 0.001$). Limited *in vitro* studies indicate that loop diuretics such as furosemide (10^{-3} M) preferentially decrease net Na^+ influx in hypertensive whites while having little

0028–0836/82/240588—02$01.00
© 1982 Macmillan Journals Ltd

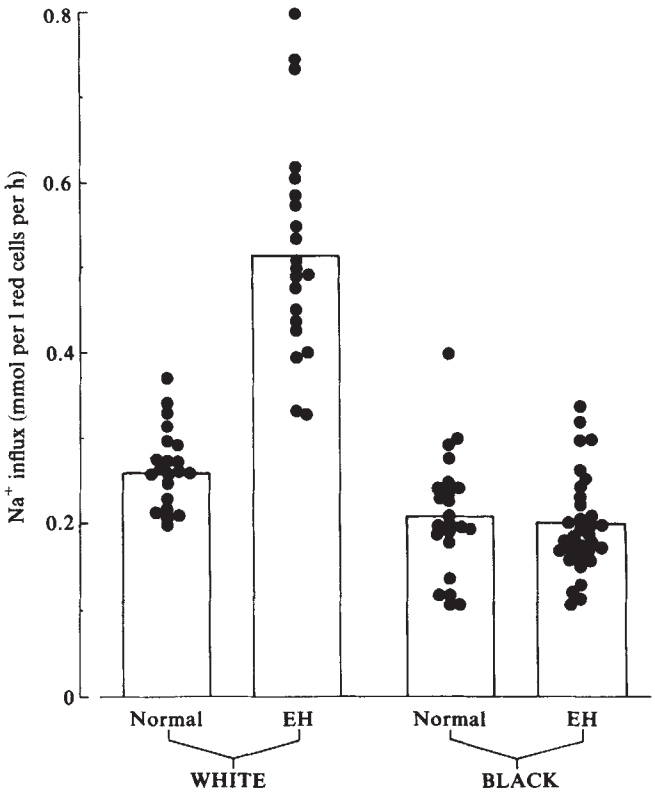


Fig. 1 Comparison of red cell Na^+ influx in normal and essential hypertensive (EH) whites and blacks. Each point represents the value for a single individual and the bars designate the mean for each group. Samples of venous blood were collected in heparinized vacutainers. Some of these were shipped on ice, with a maximum storage time of 21 h. In these cases, we included both shipment- and time-control samples from normal and hypertensive patients. The preparation of erythrocytes and assessment of $^{22}Na^+$ influx were performed as previously reported[5], except that stored samples of whole blood were re-equilibrated for 1 h at 37 °C. Red cells were washed in a buffer comprising 145 mM KCl, 5 mM NaCl, 10 mM Tris, 10 mM morpholinopropanesulphonic acid (MOPS), 10 mM D-glucose and 0.1 mM ouabain, *p*H 7.40 at 37 °C. 0.5 ml of washed, packed erythrocytes of known haematocrit (~75 vol %) was admixed with 0.5 ml of the same buffer containing 0.1 μCi of $^{22}Na^+$ (final specific activity ~20 μCi $mmol^{-1}$; NEN). The cell suspension was incubated in a shaking water bath at 37 °C for 30 min, then the cells were rapidly washed three times in 5 vols of ice-cold buffer containing 75 mM $MgCl_2$, 85 mM sucrose, 10 mM D-glucose, 10 mM Tris and 10 mM MOPS, *p*H 7.40. 0.3 ml of washed, packed cells was precipitated with 0.6 ml of 10% trichloroacetic acid. Supernatant radioactivity was assessed by liquid scintillation counting, and Na^+ influx was expressed as mmol Na^+ per l red cells per h. All determinations were made in duplicate, and the resultant values for each individual agree to within ±5%.

quantitative effect on the other three groups.

These observations suggest that either hypertension has a different genetic aetiology in the two groups or its expression (as reflected by variations in passive Na^+ transport) differs greatly. Indeed, hypertension among American blacks is also characterized by, a higher prevalence, younger age of onset and elevated risk of mortality; relatively normal renin and dopamine β-hydroxylase activities; more difficult therapeutic control; and an increased severity of secondary manifestations such as nephrosclerosis, retinopathy, accelerated atherogenesis and other vascular diseases[6-8]. The existence of a significant disparity between hypertensive whites and blacks in red cell Na^+ transport strengthens the clinical impression that hypertension in black patients differs qualitatively from that in whites. These variations in red cell Na^+ transport may be important in helping unravel the genetic basis of essential hypertension and its variable incidence and expression in different populations.

This work was supported in part by grants from the American Heart Association (81-715), the Wyoming and Montana AHA affiliates and NIH (HL-16833). N.L.E. was supported by a Faculty Research Appointment from the University of Minnesota, J.R.M. is recipient of a G. T. Evans Fellowship and J.W.E. is supported by an NIH Research Career Development Award.

Received 31 December 1981; accepted 16 April 1982.

1. Freis, E. D. *Circulation* **53**, 589–595 (1976).
2. Canessa, M., Adragna, N., Solomon, H. S., Connolly, T. M. & Tosteson, D. C. *New Engl. J. Med.* **302**, 772–776 (1980).
3. Garay, R. P., Elghozi, J.-L., Dagher, G. & Meyer, P. *New Engl. J. Med.* **302**, 769–771 (1980).
4. Meyer, P. *et al. Br. Med. J.* **282**, 1114–1117 (1981).
5. Mahoney, J. R., Etkin, N. L., McSwigan, J. D. & Eaton, J. W. *Blood* **59**, 439–442 (1982).
6. Saunders, E. & Williams, R. A. in *Textbook of Black-Related Diseases* (ed. Williams, R. A.) 333–358 (McGraw-Hill, New York, 1975).
7. Voors, A. W., Berenson, G. S., Dalferes, E. R., Webber, L. S. & Shuler, S. E. *Science* **204**, 1091–1094 (1979).
8. Warren, S. E. & O'Connor, D. R. *Am. J. Med.* **69**, 425–429 (1980).

B-cell subpopulations identified by two-colour fluorescence analysis

R. R. Hardy, K. Hayakawa, J. Haaijman & L. A. Herzenberg

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305, USA

The simultaneous and rapid measurement of the amounts of two different fluorochrome-coupled antibodies bound to single cells (two-colour immunofluorescence) provides a very powerful means for the identification of lymphocyte subpopulations[1]. Using a dual-laser fluorescence-activated cell sorter (FACS) we show that two monoclonal antibodies, anti-IgM and anti-IgD, labelled respectively with fluorescein and 'Texas red' (a new red-fluorescent dye) reveal several previously unrecognized B-cell subpopulations in mouse spleen and lymph nodes. Measured individually, these surface markers (IgM and IgD) show only that B cells are broadly heterogeneous with respect to the amount of surface immunoglobulin expressed[2]; however, measured simultaneously, they clearly define at least two B-cell subsets. One of these populations, which is predominant in spleen and constitutes the overwhelming majority of B cells in lymph nodes, is missing in CBA/N (Xid) mice known to be deficient with respect to their B-cell immune responses[3].

The B-cell populations in BALB/c spleen (Figs 1, 2) are resolved into two subsets on the basis of the amount of IgM expressed. The predominant population (labelled I in the diagram), which has relatively little surface IgM, expresses intermediate to high levels of surface IgD and constitutes ~60% of the splenic B cells (Table 1). This population is even more predominant in BALB/c lymph nodes where it constitutes >80% of the B cells.

The remaining BALB/c splenic population(s), which express 10–100 times more IgM, appear to be further subdivided according to the amount of surface IgD expressed (labelled II and III) and the organ localization of the cells. The 'high-IgD, high-IgM' population (II) constitutes ~20% of the B cells in spleen and lymph nodes. The 'low-IgD' population (III) is distinguishable more by its apparent absence from lymph nodes (see Fig. 1) than by any clear-cut demarcation in the staining pattern.

Low-angle light scatter measurements in the FACS, which provide an index of cell size[4], further distinguish these populations. Population I is the major stained component of the smaller (by scatter) cells; population II is found predominantly among the larger cells; and population III is found amongst cells in an intermediate size range. This scatter difference is another parameter which should help in isolating these subpopulations for use in functional assays.

0028–0836/82/240589—03$01.00
© 1982 Macmillan Journals Ltd

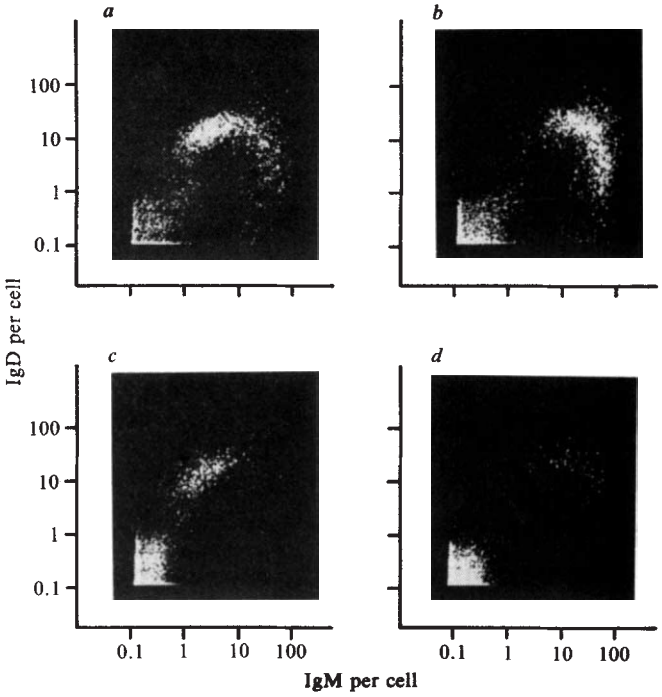


Fig. 1 B-cell subpopulations defined by two-colour FACS analysis of cells stained simultaneously with anti-IgM and anti-IgD antibodies. The individual points in the 'dot plots'[14] shown represent the amounts of IgD and IgM expressed on 3,000 individual cells from spleen and lymph nodes in CBA/N and BALB/c mice. The relative amounts of IgM and IgD per cell are indicated on the axes. *a*, BALB/c spleen; *b*, CBA/N spleen; *c*, BALB/c peripheral lymph nodes; *d*, CBA/N peripheral lymph nodes. Spleen or lymph node cells (10^6) were stained with 0.5 μg of fluorescein-labelled monoclonal rat anti-IgM[15] and 0.5 μg of biotinated monoclonal mouse anti-IgD[16] antibodies in 100 μl of biotin-free RPMI-1640 containing 10 mM HEPES buffer, 0.1% sodium azide and 3% newborn calf serum for 30 min at 0 °C. Cells were washed three times with RPMI and stained with 1 μg of Texas red–avidin in 50 μl of RPMI for 30 min at 0 °C. Cells were washed three times with RPMI, resuspended in 300 μl of the same buffer and analysed on the dual-laser FACS equipped with logarithmic amplifiers. To permit subsequent analyses, 'list-mode' data recording the scatter and two fluorescence measurements for each cell were collected on 30,000 cells using a VAX 11/780 computer.

Studies with CBA/N mice further clarify the definition of the B-cell subpopulations. These mice, which carry an X-linked B-cell defect, have almost no detectable population I cells in lymph nodes or spleen. Their splenic B cells, which consist entirely of high-IgM populations, are divided between cells comparable with the cells in populations II and III in BALB/c spleen. The CBA/N spleen also has many more IgM-, IgD-double negative cells than the BALB/c spleen. As might be expected, CBA/N spleen cells have the same scatter correlations as BALB/c spleen cells, except that the small cells lack both IgM and IgD.

There are only a few immunoglobulin-positive cells in CBA/N lymph nodes and these stain exclusively in the region of population II. Thus, when population I is absent from lymph node and consequently does not obscure the analysis of the minor B-cell populations present, populations II and III are distinguishable by their locations in the IgM, IgD fluorescence plots.

Thus the combined data from two-colour FACS analyses of CBA/N and BALB/c spleen and lymph node demonstrate that the heterogeneous expression of IgM and IgD on the whole B-cell population reflects three (or more) subpopulations, each with characteristic levels of surface immunoglobulins. Studies of the appearance of these populations during development, and the disappearance of the B-cell population(s) present predominantly during neonatal life, indicate the presence of at least one more B-cell subpopulation characterized by the lack of surface IgD (J.H., R.R.H., K.H. and L.A.H. in preparation). While it is clear that population I reaches mature level last[5] the relationship between these populations is currently unknown. An allotype-specific IgM–IgD stain carried out on sorted cells transferred to allotype congenic mice should help to clarify these relationships.

Table 1 Per cent of cells in IgM/IgD defined B-cell subpopulations

Strain	Organ	I	II	III
BALB/c	Spleen	30	10	10
	Lymph node	17	4	<1
CBA/N	Spleen	<4*	20	15
	Lymph node	<1	4	<1
CBA	Spleen	23	18	8
	Lymph node	10	4	<1

CBA/N mice lack the major B-cell subpopulation found in spleen and lymph nodes in normal mice. Subpopulations were defined by two-colour FACS analysis (Fig. 2). Data show the per cent of total spleen or lymph node cells present in the B-cell subpopulations shown in Fig. 2. Percentages for populations I and II may be in error by as much as 5% in the BALB/c and CBA determinations due to overlap between these populations.

*This value (4%) is attributable to overlap from population II.

The absence of population I in CBA/N mice fits with several previous observations: the higher IgM/IgD ratio on B cells from CBA/N spleen[6] arises from a lack of the low IgM/IgD ratio (population I) cells in these mice; the absence of certain B cell determinants [Lyb3 (ref. 7), Lyb5 (ref. 8), Lyb7 (ref. 9)] in CBA/N is apparently due to the presence of these markers exclusively on population I cells; the presence of one of these markers (Lyb3) on virtually all normal lymph node B cells, but on only a fraction of normal spleenic B cells[10], is consistent with the relative amounts of population I in the two organs.

CBA/N mice have an extensively studied immune deficiency: they do not make antibody to type II thymus-independent antigens (such as dinitrophenyl–Ficoll)[11]; their spleen cells fail to give the usual *in vitro* proliferative responses shown by normal spleen cells stimulated with anti-immunoglobulin antisera[12]; and their serum IgM and IgG3 levels are very low compared with normal mice[13]. We speculate that the B cells responding to dinitrophenyl–Ficoll or to anti-immunoglobulin antisera in the proliferative assay are contained in subpopulation I (missing in CBA/N). The B cells responsible for a large part of IgM and IgG3 antibody production also may be included in this subset. We should be able to test these hypotheses and assign functions to the B-cell populations defined here by

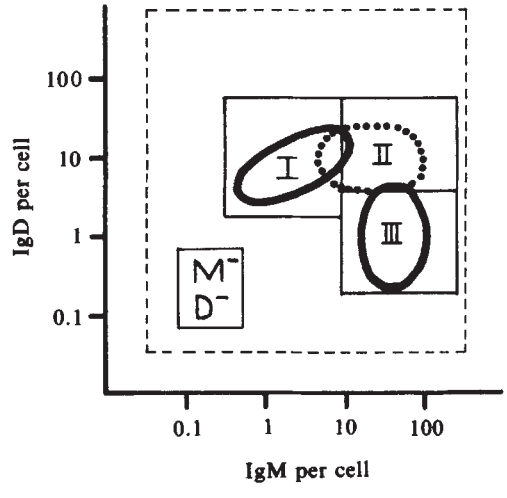


Fig. 2 Diagram of the B-cell subpopulations shown in Fig. 1 (see text). The integration bounds used to obtain the percentages in Table 1 are also shown.

© 1982 Nature Publishing Group

sorting these populations before appropriate *in vitro* and adoptive transfer assays.

We acknowledge the Institute of Experimental Gerontology, Health Organization TNO, The Netherlands, for providing CBA/N mice. We thank Dr P. Kincade for the monoclonal rat anti-IgM antibody, Dr D. Parks for invaluable assistance with the dual-laser FACS and Mr W. Moore for writing the VAX data handling programs. This work was supported in part by NIH grants CA-04681 and GM-17367.

Received 31 December 1981; accepted 15 April 1982.

1. Ledbetter, J. A., Rouse, R. V., Micklem, H. S. & Herzenberg, L. A. *J. exp. Med.* **152**, 280–295 (1980).
2. Vitetta, E. S. *et al.* in *Progress in Immunology* Vol. 3 (ed. Mandel, T.) 65–73 (North-Holland, Amsterdam, 1977).
3. Scher, I., Frantz, M. M. & Steinberg, A. D. *J. Immun.* **110**, 1396–1401 (1973).
4. Fathman, C. G., Small, M., Herzenberg, L. A. & Weissman, I. L. *Cell. Immun.* **15**, 109–128 (1975).
5. Hardy, R. R., Hayakawa, K. & Herzenberg, L. A. *Ann. N.Y. Acad. Sci.* (in the press).
6. Finkelman, F. D., Smith, A. H., Scher, I. & Paul, W. E. *J. exp. Med.* **142**, 1316–1321 (1975).
7. Huber, B., Gershon, R. K. & Cantor, H. *J. exp. Med.* **145**, 10–20 (1977).
8. Ahmed, A. *et al. J. exp. Med.* **145**, 101–110 (1977).
9. Subbarao, B. *et al. J. Immun.* **122**, 2279–2285 (1979).
10. Huber, B., Gathings, W. E. & Cooper, M. D. in *B Lymphocytes in the Immune Response* (eds Cooper, M. D., Mosier, D. E., Scher, I. & Vitetta, E.) 125–130 (Elsevier, Amsterdam, 1979).
11. Mosier, D. E., Mond, J. J. & Goldings, E. A. *J. Immun.* **119**, 1874–1878 (1977).
12. McKenzie, I. F. C. & Potter, T. *Adv. Immun.* **27**, 241 (1979).
13. Perlmutter, R. M. *et al. J. exp Med.* **149**, 993–998 (1979).
14. Herzenberg, L. A. & Herzenberg, L. A. in *Handbook of Experimental Immunology* 3rd edn (ed. Weir, D. M.) 22.1–22.21 (Blackwell Scientific, Oxford, 1978).
15. Kincade, P. W., Lee, G., Sun, L. & Watanabe, T. *J. Immun. Meth.* **42**, 17–26 (1981).
16. Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **8**, 115–129 (1978).

Target antigens of purified human immunoglobulins which inhibit growth of *Plasmodium falciparum in vitro*

Graham V. Brown*, Robin F. Anders*, Graham F. Mitchell* & Peter F. Heywood†

* Laboratory of Immunoparasitology, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria 3050, Australia
† The Papua New Guinea Institute of Medical Research, Madang, Papua New Guinea

A prerequisite for an antigenically defined vaccine against malaria is the identification of parasite antigens which are necessary or sufficient for the induction and expression of host-protective immunity. As malaria infection in man can be treated by passive transfer of serum or immunoglobulin from adults in an endemic area[1] it is reasonable to propose that sera from clinically defined patient groups and functionally defined immunoglobulin may be used as probes to identify the relevant parasite antigens involved in host protection ('host-protective antigens')[2]. We have now purified immunoglobulin from a large number of individuals living in the endemic area, tested for inhibitory effects on *in vitro* growth of *Plasmodium falciparum* originating from the same area, then examined antibody specificities by immunoprecipitation analysis with biosynthetically labelled proteins of blood stages of the parasite. Two-dimensional gel analysis of immunoprecipitates revealed a correlation between inhibition of parasite growth and increased antibody binding to two proteins of molecular weight (M_r) ~96,000.

Serum was obtained with informed consent from healthy adult blood donors living in the Madang region of Papua New Guinea (PNG), where malaria is endemic, at a time of the year when transmission was low. Purified immunoglobulin from each donor (prepared using protein A–Sepharose, Pharmacia) was screened for its ability to inhibit the *in vitro* growth of a PNG isolate of *P. falciparum* by a modification of a method described previously[3]. In all experiments we used a single parasite isolate which has been maintained in our laboratory since it was obtained from an infected child in Madang in 1979. Briefly, immunoglobulin was dialysed against culture medium (RPMI 1640 with sodium bicarbonate and HEPES buffer) for 36 h and adjusted to a final concentration of 5 mg ml^{-1} in culture medium. After addition of normal human serum to a final concentration of 10%, the 'immunoglobulin-supplemented medium' was assayed for its ability to inhibit the growth of *P. falciparum in vitro*. In this assay, 200 μl of parasitized blood group O erythrocytes in culture medium (8% v/v) were added to a series of flat-bottomed wells of a microtitre tray and cultured in a candle jar using the method devised by Trager and Jensen[4]. Each day, culture medium was changed (using normal or test medium) until day 4, when parasitaemia was assessed by uptake of ^{3}H-leucine into trichloracetic acid-precipitable material, such uptake having been shown to correlate with the percentage of parasitized cells (G.V.B., unpublished). Using this assay, immunoglobulin prepared from 4 of 39 sera inhibited growth by more than 50%.

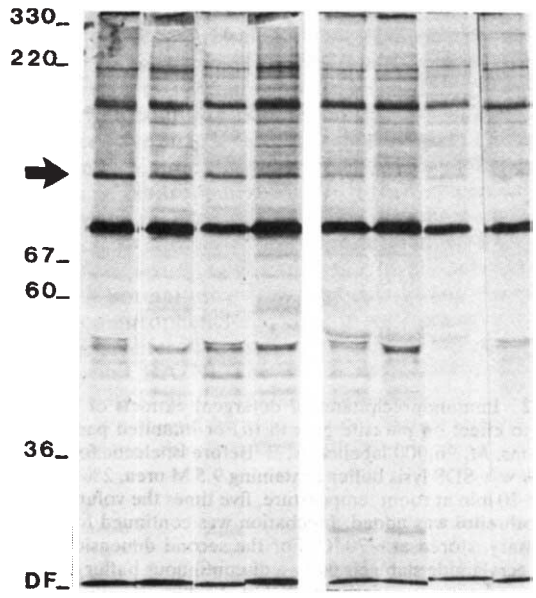


Fig. 1 SDS-polyacrylamide gel electrophoresis of immunoprecipitates of detergent extracts of ^{35}S-methionine-labelled *P. falciparum* antigen with protein A-purified immunoglobulin from adults living in the endemic area. The arrow (→) indicates a band which is dominant in immunoprecipitates with the four most inhibitory immunoglobulins (left-hand lanes), compared with four non-inhibitory immunoglobulins (right-hand lanes). Molecular weight markers are given (×10^{-3}). DF, dye front.

Pools were made of five sera containing immunoglobulin which inhibited parasite growth and five sera containing immunoglobulin which, at the same concentration (5 mg ml^{-1}), had no inhibitory effect on parasite growth. IgG was prepared from them using a different method from that used for the individual samples, by precipitation with 45% ammonium sulphate followed by ion exchange chromatography on DEAE-cellulose. The IgG was then dialysed against culture medium and assayed for inhibition of *P. falciparum* growth *in vitro*. The pooled inhibitory IgG caused at least 65% inhibition of parasite growth compared with non-inhibitory IgG (50,800 ± 420 c.p.m. for non-inhibitory IgG compared with 17,150 ± 750 c.p.m. for inhibitory IgG using the radioisotopic readout for parasite growth inhibition).

Biosynthetically labelled parasite antigens were prepared by culturing the same isolate of PNG *P. falciparum* as used for the inhibition studies for 14 h in methionine-free medium (RPMI 1640 Select-Amine Kit, Gibco) supplemented with ~200 μCi ml^{-1} ^{35}S-methionine (800 Ci mmol^{-1}, Radiochemical Centre). A detergent extract of parasitized cells was made using 0.5% Triton X-100 in NaCl, EDTA and Tris buffer[5] (Triton NET) containing methionine (2 mg ml^{-1}). After centrifugation at 25,000*g* for 30 min, the supernatant was passed over a Sephadex G-25M column (PD10 column, Pharmacia) and eluted in 0.5% Triton NET supplemented with methionine.

0028–0836/82/240591—03$01.00
© 1982 Macmillan Journals Ltd

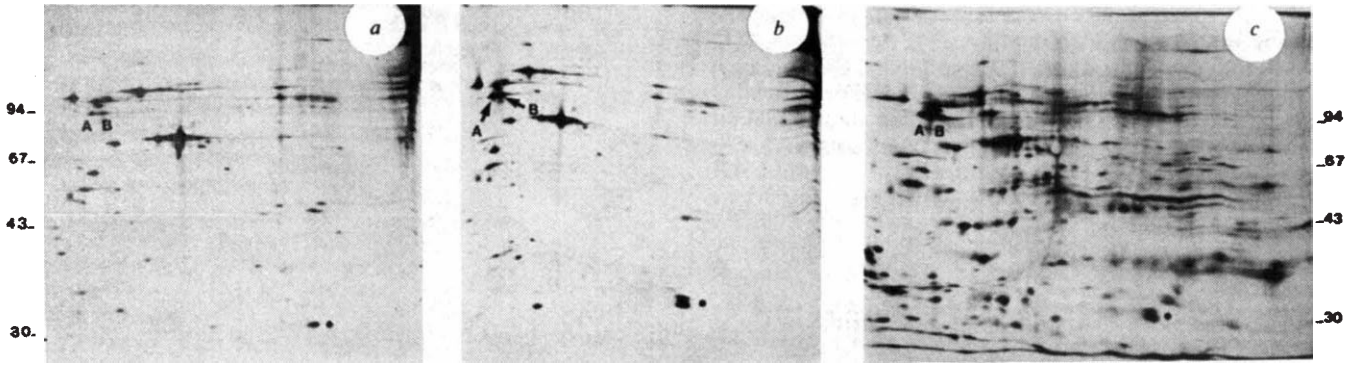


Fig. 2 Immunoprecipitates of detergent extracts of ^{35}S-methionine-labelled *P. falciparum* (*c*) by purified human immunoglobulin which had no effect on parasite growth (*a*) or inhibited parasite growth (*b*). Inhibitory IgG contained a high titre of antibody to the two acidic proteins, M_r 96,000 labelled A, B. Before isoelectic focusing, the antigen–antibody complexes bound to *S. aureus* protein A were solubilized in 2% w/v SDS lysis buffer containing 9.5 M urea, 2% ampholines (4 : 1 v/v *p*H 5–8 : *p*H 3.5–10 ampholines, LKB) and 50 mM dithiothreitol. After 20 min at room temperature, five times the volume of lysis buffer containing 2% Triton X-100, 2% ampholines, 9.5 M urea and 50 mM dithiothreitol was added. Incubation was continued for 15 min at room temperature before samples were centrifuged before analysis, or if necessary, stored at −70 °C. For the second dimension, electrophoresis in reducing conditions in the presence of SDS was performed on 10% acrylamide slab gels using a discontinuous buffer system[6]. Gels were prepared for fluorography using EN^{3}HANCE (NEN) before drying down and exposure to Kodak X-Omat S film. Numbers refer to the molecular weights ($\times 10^{-3}$) of M_r standards (Pharmacia Fine Chemicals Low Molecular Weight Calibration Kit) which were run on each slab and identified by Coomassie blue staining. Acidic region is to the left of each gel.

IgG isolated from each pool of sera, or protein A-purified immunoglobulin from individual sera, were reacted with biosynthetically labelled parasite proteins and antigen–antibody complexes precipitated with *Staphylococcus aureus* as described previously[3] before analysis by one[6] and two-dimensional gel electrophoresis[7]. Approximately equal amounts of radioactivity were precipitated from preparations of solubilized biosynthetically labelled infected cells by IgG from the two pools of sera. Examination of the fluorographs (Fig. 1) showed that the majority of proteins precipitated by non-inhibitory immunoglobulins (right-hand lanes) were precipitated to the same extent by the inhibitory immunoglobulins (left-hand lanes). However, one band of M_r 96,000 (indicated by the arrow) was dominant in immunoprecipitates using inhibitory immunoglobulins but was poorly precipitated by non-inhibitory immunoglobulins. Two-dimensional analysis of immunoprecipitates using IgG from sera pools confirmed that there were many similarities in proteins precipitated by non-inhibitory IgG (Fig. 2*a*) and inhibitory IgG (Fig. 2*b*). However, as in the one-dimensional analysis there was a difference in the two-dimensional gel patterns at M_r 96,000 in that two acidic proteins (labelled A, B in Fig. 2) were much more prominent in precipitates with inhibitory IgG. Different antigen preparations made from the same isolate of *P. falciparum* contained varying proportions of different parasite proteins[8], presumably because the proportion of parasites at each stage in the life cycle in asynchronous culture varied each time an antigen preparation was made. For example, other antigen preparations of the same isolate (not shown) had a greater proportion of high molecular weight proteins. Antigen A (Fig. 2) was invariably well represented while antigen B was sometimes present in only small amounts. The representation of several minor proteins also differed between the immunoprecipitates, including a protein with approximately the same isoelectric point (*p*I) as the M_r 96,000 proteins but of M_r ~40,000. Another difference in precipitation of a low molecular weight protein (* in Fig. 2) has not been emphasized because it was not reproducible.

We thus conclude that natural infection in man leads to antibody production which can inhibit the growth of *P. falciparum in vitro*. We found that only 4 of 39 immunoglobulin preparations inhibited parasite growth by >50% while some samples (9 of 39) actually increased growth beyond control levels (a result consistent with previous studies using unfractionated sera[9]). Thus, it seems that the effects of immunoglobulin on parasite growth *in vitro* may be a balance between growth promoting and inhibiting activities. In the present study, the population donating sera at a time of the year when transmission was low contained relatively few individuals whose serum had net inhibitory activity.

IgG which inhibited parasite growth immunoprecipitated two proteins (M_r 96,000) more efficiently than did IgG with no inhibitory effect on parasite growth. In previous studies using one-dimensional gel analysis to compare immunoprecipitates of inhibitory and non-inhibitory serum[3], we noted differences in recognition by the sera of a protein band of approximately the same molecular weight. Thus we have identified at least two proteins (designated A and B), antibody to which correlates with *in vitro* inhibition of growth of *P. falciparum*. The observations strongly suggest that these protein antigens may be important in the induction (and expression) of host-protective immunity in man as a result of exposure to infection in an endemic area for malaria. We have examined the antigenic composition of 15 other isolates of *P. falciparum* from Papua New Guinea and all contain antigen A as a major biosynthetically labelled protein.

Other investigators have examined the effect of human[9] or immune monkey[10–12] sera on the growth of *P. falciparum in vitro* but have been unable to correlate the inhibition with antibody to specific antigens. Hybridoma-derived murine monoclonal antibodies which inhibit the growth of *P. falciparum in vitro* have been reported recently[13]. One of these hybridoma antibodies reacted with two polypeptides of M_r 96,000 and 36,000 (ref. 13). Our finding that purified human IgG which inhibits parasite growth *in vitro* has excess antibody directed against two proteins of M_r 96,000 suggests that the hybridoma-derived antibody has the same specificity as that of antibodies contained in human IgG preparations which block parasite growth.

Additional proteins must be present in the blood stages of *P. falciparum* capable of inducing a host protective immune response in man which have not been demonstrated by the techniques we have outlined. Some immunoglobulin preparations causing 30–50% inhibition of parasite growth *in vitro* did not precipitate relatively large amounts of antigens A and B, confirming that other specificities are also important. Our techniques would not detect proteins with little methionine, proteins not solubilized by detergent extraction or after immune complex precipitation, very basic proteins, or non-protein antigens. Note also that other correlates with protection may be important at times of high transmission with multiple antigenic variants, high parasite loads and circulating immune complexes, and with the

© 1982 Nature Publishing Group

possibility of rise and fall of competing facilitatory and inhibitory antibodies. Furthermore, analysis of other mechanisms in which immunoglobulin may have a host-protective role, such as opsonization, may reveal yet other antigenic correlates. It remains to be determined whether the proteins identified by this comparative immunoprecipitation approach are immunogenic after isolation, and whether they induce host protection or inhibitory antibodies in appropriate model systems.

The work was supported by the National Health and Medical Research Council of Australia, the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the Rockefeller Foundation Great Neglected Diseases of Mankind Network. We acknowledge the technical assistance of Raelene Grumont.

Received 23 November 1981; accepted 14 April 1982.

1. Cohen, S., McGregor, I. A. & Carrington, S. C. *Nature* **192**, 733–737 (1961).
2. Knopf, P. M., Brown, G. V., Howard, R. J. & Mitchell, G. F. *Aust. J. exp. Biol. med. Sci.* **57**, 603–615 (1979).
3. Brown, G. V., Anders, R. F., Stace, J. D., Alpers, M. P. & Mitchell, G. F. *Parasite Immun.* **3**, 283–298 (1982).
4. Jensen, J. B. & Trager, W. *J. Parasit.* **63**, 883–886 (1977).
5. Kessler, S. W. *J. Immun.* **115**, 1482–1490 (1975).
6. Laemmli, U. K. & Favre, M. *J. molec. Biol.* **80**, 575–599 (1973).
7. O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
8. Brown, G. V., Coppel, R. L., Vrbova, H., Grumont, R. J. & Anders, R. F. *Exp. Parasit.* **53**, 279–284 (1982).
9. Wilson, R. J. M. & Phillips, R. S. *Nature* **263**, 132–134 (1976).
10. Reese, R. T. & Motyl, M. R. *J. Immun.* **123**, 1894–1899 (1979).
11. Butcher, G. A. & Cohen, S. *Immunology* **23**, 503–519 (1972).
12. Campbell, G. H., Mrema, J. E. K., O'Leary, T. R., Jost, R. C. & Rieckman, K. H. *Bull. Wld Hlth Org.* **57**, Suppl. 1, 219–215 (1979).
13. Perrin, L. H., Ramirez, E., Lambert, P. H. & Miescher, P. A. *Nature* **289**, 301–303 (1981).

Cancer metastasis is selective or random depending on the parent tumour population

J. E. Talmadge & I. J. Fidler

Cancer Metastasis and Treatment Laboratory,
NCI–Frederick Cancer Research Facility, PO Box B, Frederick,
Maryland 21701, USA

The issue of whether metastases result from the random survival of cells released from the primary tumour or from the selective growth of specialized subpopulations of cells having properties that allow them to complete the metastatic process is important to our understanding of tumour biology[1,2]. Previous studies have provided indirect evidence that the process of metastasis favours the survival of cells having a metastatic phenotype(s)[3–8]. Direct evidence that the process is selective would be provided by the demonstration that cells populating spontaneous metastases are, in general, more metastatic than most of the parent tumour cells; some[9–11], but not all[12,13], previous reports have failed to provide such evidence, suggesting that the process is random. These discrepancies could be due to differences in the biological characteristics of the various tumour systems studied and in the experimental conditions[14]. In the present study, these variables were minimized by using three metastatic variant cell lines of the B16 melanoma. We report that, even in controlled conditions, the process of metastasis can appear as either selective or random depending on the nature of the initial population. Specifically, metastasis by the unselected, poorly metastatic parent B16 tumour was indeed selective. In contrast, metastasis by previously selected B16 lines appeared to be random.

To determine whether metastases contain cells having increased metastatic potential, we used three metastatic tumour cell variants isolated from the B16 melanoma. B16-F1 is an unselected tumour cell line that metastasizes poorly when tumour cells are implanted intravenously (i.v.; experimental metastasis) or subcutaneously (s.c.; spontaneous metastasis)[15]. The B16-F10 cell line[15] was selected for its ability to colonize the lung after i.v. administration, and the B16-BL6 cell line[16] was selected *in vitro* for its invasiveness, demonstrated by a high incidence of spontaneous metastases after s.c. implantation. All three tumour cell lines were implanted into the footpads of syngeneic 6-week-old C57BL/6 female mice (20 per group). After 5–7 weeks, when one or two mice from each group became listless, the entire group was killed. Only 5 of 20 mice injected with B16-F1 cells exhibited lung metastases (median number of lung colonies 0, range 0–12); 8 of 20 mice injected with B16-F10 cells developed metastases (median 10, range 0–60); 18 of 20 mice injected with B16-BL6 cells exhibited gross metastases (median 50, range 0–200). In all groups, well-isolated pulmonary metastases were surgically excised and established in culture as individual cell lines. The metastatic potential of cells from the parent tumour variants and several of their respective spontaneous metastases was determined in assays measuring experimental and/or spontaneous metastasis. A recent report suggested that the resection of a 'primary' B16 tumour could lead to the production of metastases[12], therefore we divided the mice bearing s.c. tumours into two treatment groups. In the first group, the leg carrying a 1.2–1.5 cm diameter tumour was amputated at mid-femur, thus including the popliteal lymph node, whereas in the second group of mice, the tumour was not resected.

Table 1 Ability of cells from B16 melanoma variants and their spontaneous metastases to form lung tumour colonies

Tumour line	No. of cells injected ($\times 10^3$)	Median no. of pulmonary foci (range)	*P*
B16-F1	50	5 (2–35)	—
M1	50	75 (30–>300)	0.003
M2	50	61 (29–103)	0.02
M3	50	91 (0–237)	0.002
M4	50	127 (41–>300)	0.0002
B16-BL6	50	48 (19–237)	—
M1	50	>300 (80–>300)	0.002
M2	50	170 (1–262)	0.03
M3	50	>300 (170–>300)	0.001
M4	50	>300 (172–>300)	0.0002
B16-F10	10	11 (1–19)	—
M1	10	11 (6–38)	0.86
M2	10	21 (3–42)	0.08
M3	10	6 (3–25)	0.47
M4	10	17 (4–35)	0.078
B16-F10	50	152 (22–219)	—
M1	50	217 (97–297)	0.12
M2	50	251 (1–300)	0.02
M3	50	130 (41–156)	0.3
M4	50	183 (93–>300)	0.96

Spontaneous metastases (M1–M4) from each of the tumours were obtained from mice bearing footpad tumours. Cells were collected from mid-log phase cultures by 1 min treatment with 0.25% trypsin, 0.02% EDTA solution. Single-cell suspensions of >95% viability were adjusted to the desired concentrations in Hanks' balanced salt solution ($Ca^{2+}+Mg^{2+}$-free) and 0.2 ml was injected into the lateral tail vein. Mice were necropsied 21 days later, and the number of lung tumour colonies was determined. Simultaneous comparisons of treatment versus control were based on the Kruskal–Wallis rank sums using the Miller approximation procedure[21]. For the analysis, >300 lung colonies were treated as being equal to 300. One-tailed *P* value is shown; 10 mice per group.

The data in Table 1 demonstrate that cells collected from spontaneous metastases from the poorly metastatic B16-F1 line produced significantly (15–25 times) more lung tumour colonies after i.v. implantation than an equal number of cells from the parent line. With the previously selected B16-BL6 line[16], cells collected from spontaneous metastases produced only 3–6 times more lung tumour colonies than the parent line. The B16-F10 cell line, which was specifically selected for the ability to produce pulmonary colonies[15], contrasted even more sharply with the B16-F1 cell line. Spontaneous metastases of B16-F10 tumours did not exhibit any increased propensity for experimental metastasis, regardless of whether the inoculum contained a high or low number of cells.

In the assays of spontaneous metastasis, individual variability in metastatic potential of cells populating different spontaneous metastases can be expected[2]. Thus, we analysed pooled data of metastases versus their respective parental tumours (Table 2). Once again, cells collected from spontaneous metastases originating from the B16-F1 tumour were significantly more metastatic than those originating from the parent tumour ($P=0.007$). Cells collected from spontaneous metastases of B16-

0028–0836/82/240593—02$01.00

© 1982 Macmillan Journals Ltd

Table 2 Production of spontaneous metastases by cells obtained from parent B16 melanoma variants and their spontaneous metastases

Tumour line	Incidence of spontaneous metastasis in mice with: Primary tumour resected	P	Primary tumour not resected	P	Total	P
B16-F1	3/7		2/10		5/17	
M1	4/7		6/10		10/17	
M2	7/7	0.025	7/11	0.009	14/18	0.007
M3	8/8		6/11		14/19	
M4	8/9		8/8		16/17	
B16-F10	6/8		2/8		8/16	
M1	8/10		5/10		13/20	
M2	9/10	0.34	7/10	0.05	16/20	0.0426
M3	8/9		7/10		15/19	
M4	9/10		6/9		15/19	
B16-BL6	8/8		6/8		14/16	
M1	8/8		4/8		12/16	
M3	10/10	0.64	5/10	0.31	15/20	0.2165
M3	6/6		6/8		12/14	
M4	6/8		7/12		13/20	

Primary tumours were initiated intramuscularly in 0.05 ml Hanks' balanced salt solution ($Ca^{2+}+Mg^{2+}$-free) into a posterior footpad. The tumour-bearing leg and popliteal lymph node were resected from 10 mice when the tumours reached 1.2–1.5 cm in diameter. Some mice died immediately after the resection and were thus excluded from the study. Morbid mice were necropsied, and the number of metastases was determined. The metastatic ability of the parent tumour was compared with that of cells recovered from metastases by the Fisher exact test using the data for resected or non-resected subcutaneous tumours as well as the combined data[21]. Combining the data from the two treatment groups increased sample size and yielded a more sensitive test.

F10 tumours exhibited an increased propensity for spontaneous metastasis ($P=0.0426$) but, as shown in Table 1, not for experimental metastasis. Finally, cells from spontaneous metastases produced by B16-BL6, a tumour that exhibits a very high incidence of spontaneous metastasis (>80%), did not differ significantly ($P=0.2165$) from the parent cells in metastatic potential.

Our findings demonstrate that metastases produced from the unselected, heterogeneous, and poorly metastatic B16-F1 cell line indeed contained cells having enhanced metastatic potential. In contrast, when the parent population was previously selected and was already performing at near peak ability (B16-F10 for lung colonization or B16-BL6 for spontaneous metastasis), further selection for the particular phenotype could not be achieved. These results clearly support the concept that metastasis is a selective process and agree with the recent demonstrations[13,17] that non-selective processes (adaptation) do not result in the outgrowth of highly metastatic cells.

We have shown that the B16 melanoma, which originated in 1954, rapidly changes its characteristics depending on the *in vivo* conditions in which it is maintained[18]. For this reason the B16 tumour maintained in one laboratory may not be exactly equivalent to a B16 tumour propagated in another laboratory. Specifically, Trope[19] proposed that the repeated trocar or fragment passage of heterogeneous tumours leads to loss of their diversity and imposition of uniformity on the transplanted tumours. Recent studies have substantiated this hypothesis and have demonstrated that repeated passages with tumour fragments appear only to exacerbate this artefact[20]. Clearly, to study selection during the process of metastasis, it is critical that the starting tumour be heterogeneous. It is difficult to envisage how successful selection could occur with a tumour composed of a uniform cell population. Furthermore, as numerous artefacts can alter the outcome of metastatic assays[14], absolute standardization of experimental conditions is necessary if the results of such assays are to be reproducible. In any event, we conclude from our studies that in controlled, experimental conditions, metastases produced by the heterogeneous B16-F1 melanoma cell line have a specialized subpopulation of cells with enhanced metastatic properties, supporting the conclusion that metastasis is a selective process.

This research was sponsored by the NCI under contract No. NO1-CO-75380 with Litton Bionetics, Inc.

Received 9 December 1981; accepted 6 April 1982.

1. Fidler, I. J. & Kripke, M. L. *Science* **197**, 893–895 (1977).
2. Poste, G. & Fidler, I. J. *Nature* **283**, 139–146 (1980).
3. Nicolson, G. L. & Winkelhake, J. L. *Nature* **255**, 230–232 (1975).
4. Briles, E. B. & Kornfeld, S. *J. natn. Cancer Inst.* **60**, 1217–1222 (1978).
5. Brunson, K. W., Beattie, G. & Nicolson, G. L. *Nature* **272**, 543–545 (1978).
6. Tao, T. W. & Burger, M. M. *Nature* **270**, 437–438 (1977).
7. Talmadge, J. E., Meyers, K. M., Prieur, D. J. & Starkey, J. R. *Nature* **284**, 622–624 (1980).
8. Raz, A. & Hart, I. R. *Br. J. Cancer* **42**, 331–341 (1980).
9. Giavazzi, R., Alessandri, G., Spreafico, F., Garattini, S. & Mantovani, A. *Br. J. Cancer* **42**, 464–472 (1980).
10. Mantovani, A., Giavazzi, R., Alessandri, G., Spreafico, F. & Garattini, S. *Eur. J. Cancer* **17**, 71–76 (1981).
11. Eccles, S. A., Heckford, S. E. & Alexander, P. *Br. J. Cancer* **42**, 252–259 (1980).
12. Stackpole, C. W. *Nature* **289**, 798–800 (1981).
13. Raz, A., Hanna, N. & Fidler, I. J. *J. natn. Cancer Inst.* **66**, 183–189 (1981).
14. Fidler, I. J. *Meth. Cancer Res.* **15**, 399–439 (1978).
15. Fidler, I. J. *Nature new Biol.* **242**, 148–149 (1973).
16. Hart, I. R. *Am. J. Path.* **97**, 587–600 (1979).
17. Nicolson, G. L. & Custead, S. E. *Science* **215**, 176–178 (1982).
18. Fidler, I. J. & Nicolson, G. L. in *Cancer Biology Reviews* (eds Marchalonis, J., Hanna, M. G. Jr & Fidler, I. J.) 171–234 (Marcel-Dekker, New York, 1981).
19. Trope, C. in *Design of Models for Testing Cancer Therapeutic Agents* (eds Fidler, I. J. & White, R. J.) 62–79 (Van Nostrand Reinhold, New York, 1981).
20. Fidler, I. J. & Hart, I. R. *Cancer Res.* **41**, 3266–3267 (1981).
21. Miller, R. G. Jr *Simultaneous Statistical Inference*, 6–10, 167 (McGraw-Hill, New York, 1966).

Selective killing of malignant cells in a leukaemic rat bone marrow using an antibody–ricin conjugate

P. E. Thorpe*, D. W. Mason†, A. N. F. Brown*, S. J. Simmonds†, W. C. J. Ross‡, A. J. Cumber‡ & J. A. Forrester‡

* Imperial Cancer Research Fund, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK
† MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK
‡ Chester Beatty Research Institute, Fulham Road, London SW3 6JB, UK

Graft-versus-host disease and marrow graft rejection are two major problems in the treatment of leukaemia with whole body irradiation, high-dose chemotherapy and bone marrow transplantation. Both these problems could be avoided if the patient's own bone marrow could be rendered aleukaemic *in vitro* and then reinjected into the patient after the radiochemotherapy had been completed. Here we describe a model system in which lethally irradiated rats were injected with a mixture of bone marrow cells and leukaemic cells and in which the development of leukaemia in the recipient rats was prevented by incubating the cells, before injection, with a monoclonal antibody–ricin conjugate that was selectively toxic to the leukaemic cells. Nonspecific toxicity of the conjugate to haematopoietic stem cells was blocked by carrying out the incubation in the presence of lactose which competitively antagonized the binding of the ricin moiety of the conjugate to galactose residues on the cell surface[1–3].

The rat leukaemia used in these experiments is of T-cell origin and developed originally in a PVG rat that had received two injections of radioactive tungsten trioxide ($^{185}WO_3$) into its lymphoid tissue approximately 26 weeks earlier[4]. Injection of graded doses of leukaemic cells into fresh PVG hosts showed that a 1,000-fold reduction in cell dose delayed the onset of the leukaemia by 8-10 days, indicating a cell cycle time of a little less than 24 h. The injection of 100 leukaemic cells into syngeneic recipients was always fatal and of two rats given approximately 10 cells, one developed leukaemia and the other did not.

The derivation of the monoclonal antibody, W3/25, with which the ricin was coupled is given in ref. 5. It is an IgGl class antibody and reacts with high affinity[5] with rat thymocytes[6], T

 © 1982 Macmillan Journals Ltd

helper cells[7,8] and macrophages[9] but not with bone marrow haematopoietic stem cells (D.W.M., unpublished). Although the leukaemic cells express the W3/25 antigen (B. Roser, personal communication), flow cytofluorograph analysis indicates that the level of antigen expression varies widely between cells. Attempts to produce a more homogeneous population by *in vivo* cloning have been unsuccessful.

Initially, we attempted to prepare a specific cytotoxic agent for the leukaemic cells by coupling W3/25 antibody by a disulphide bond directly to the ribosome-damaging A-subunit of ricin[10]. This approach was used successfully by Krolick *et al.*[11], who recently described a conjugate of rabbit anti-immunoglobulin antibody and ricin A-chain which was capable of killing leukaemic cells in infiltrated murine bone marrow *in vitro* without destroying the haematopoietic stem cells. The W3/25–ricin A-chain conjugate, however, had almost no cytotoxic action on the leukaemic cells or on rat splenic T lymphocytes[10], even when used at concentrations which saturated the cell-surface antigens. The difference in the cytotoxic performance of the two conjugates typifies the variable effectiveness of A-chain conjugates prepared in several laboratories (reviewed in refs 10, 12, 13) and suggests that only certain types of cell-surface antigen can mediate the translocation of the A-chain into the cytosol. Therefore, for the present work, we linked intact ricin to W3/25 antibody, thereby retaining any properties of the B-chain which are necessary to effect A-chain penetration. The conjugate was prepared using the *N*-hydroxysuccinimidyl ester of chlorambucil as described previously[10] and was used in the presence of lactose, which antagonized the binding of the ricin B-chain moiety to galactose residues on cells lacking the W3/25 antigen.

The W3/25–ricin conjugate was an extremely effective inhibitor of protein synthesis by PVG leukaemic cells in tissue culture, as shown by their markedly reduced capacity to incorporate ^{3}H-leucine. Leucine uptake was reduced by 50% on treatment of the cells for 1 h at 37 °C with the conjugate at a concentration of 1.4 ng ml^{-1} ricin (2.2×10^{-11} M). Even in the absence of lactose the W3/25–ricin conjugate was about 10-fold more potent than native ricin, and about 100-fold more potent than a control conjugate made with another monoclonal antibody, MRC OX8, which does not bind to the leukaemic cells (Fig. 1*a*). W3/25 antibody alone, at concentrations as high as 10 μg ml^{-1}, did not affect the rate of protein synthesis of the cells. When 100 mM lactose was included in the cultures, it strongly inhibited the nonspecific toxic actions of native ricin and of MRC OX8–ricin but was virtually without effect on the toxicity of the W3/25–ricin conjugate (Fig. 1*b*).

The percentage of cells which survived treatment with the W3/25–ricin conjugate *in vitro* was estimated by comparing the time that elapsed between the injection of the treated cells into PVG recipients and the appearance of leukaemic cells in the peripheral blood with that observed when graded doses of untreated leukaemic cells were injected. Rats which received 10^7 leukaemic cells that had been treated with the conjugate at 2.1 μg ml^{-1} ricin (10 μg ml^{-1} IgG) for 1 h at 37 °C in 100 mM lactose and then washed twice in lactose-containing medium developed leukaemia 8–10 days later than the recipients of untreated cells, indicating that the conjugate had destroyed about 99.9% of the leukaemic cells. Neither the treatment of cells in the presence of lactose with W3/25 antibody alone at 20 μg ml^{-1} nor treatment with MRC OX8–ricin at 4.3 μg ml^{-1} ricin delayed the appearance of leukaemia in the recipients. The leukaemic cells which grew in the rats that had received cells treated with the W3/25–ricin conjugate were as heterogeneous in expression of the W3/25 antigen as the parent population, and their treatment with the conjugate before transfer to further recipient rats again resulted in the destruction of at least 99.9% of the cells. Thus, the leukaemic cells which survived exposure to the conjugate do not seem to be a minor subpopulation which is insensitive to the conjugate or lacking in the W3/25 antigen. It is possible, however, that the cells which survived were at a stage in their cell cycle at which the W3/25 antigen was only weakly expressed.

To discover whether leukaemic cells in bone marrow could be killed by the antibody–toxin conjugate without destroying the haematopoietic stem cell activity, a mixture of such cells was incubated with the conjugate and, after washing, injected into PVG hosts which had been sublethally irradiated with 650 rad. The bone marrow cells were obtained from rats that were congeneic with PVG but whose immunoglobulin light-chain genes were derived from the DA strain and were therefore of the *1-a* allotype rather than the *1-b* allotype of PVG rats. Survival of stem cells in the injected mixture was assessed from their ability to compete with the stem cells in the irradiated

Table 1 In mixtures of leukaemic cells and bone marrow cells the malignant cells are specifically killed by the antibody–ricin conjugate

Expt	Cells injected	Incubation conditions	No. of recipients	Day of appearance of leukaemia	% B-cell chimaerism
1	10^3 Leuk. + 10^7 BM	Med.	3	16, 17, 18	—
		Med. + conj.	3	>67, >67, >67	34.0, 37.5, 38.6
	10^4 Leuk. + 10^7 BM	Med.	3	16, 16, 16	—
		Med. + conj.	3	36, >67, >67	—, 30.6, 46.5
	10^7 BM only	Med.	3	—	61.4, 66.2, 67.8
2	10^3 Leuk. + 10^7 BM	Med.	3	19, 19, 20	—
		Med. + conj.	3	—*, —*, >73	28.7
	10^4 Leuk. + 10^7 BM	Med.	3	12, 16, 17	—
		Med. + conj.	3	23, >73, >73	—, 27.8, 29.5
	10^7 BM only	Med.	3	—	52.2, 54.3, 55.0

Med., medium: Dulbecco (A+B) (phosphate-buffered saline + $CaCl_2$ and $MgCl_2$) + 100 mM lactose. Med. + conj., medium + conjugate: as for 'Med.' except that W3/25–ricin conjugate containing 2.1 μg ml^{-1} ricin and 10 μg ml^{-1} antibody was added. The conjugate had a molecular weight of 350,000 and contained antibody and ricin in a molar ratio of 2.0:1, suggesting that it comprised two molecules of antibody and one of toxin. It was more effective at destroying leukaemic cells in bone marrow than a simple conjugate (molecular weight 210,000) containing one molecule each of antibody and toxin. Recipient PVG rats were matched for age and sex. Rats in expt 1 were given 650 rad ^{137}Cs irradiation before cell transfer. Rats in expt 2 received 1,000 rad. The day of appearance of leukaemia is defined as the day at which the white blood cell count reached 2×10^4 mm^{-3}. All rats that reached this level subsequently became severely leukaemic, with a white blood cell count $>10^5$ mm^{-3}, at which time they were killed by ether anaesthesia. B-cell chimaerism was determined 39 days after cell injections for expt 1 and on day 32 for expt 2. Cervical lymph node cells were assayed for B-cell chimaerism by comparing the percentage of cells labelled by PVG rat F(ab')$_2$ anti-rat *1-a* allotype antibody with that labelled by rabbit F(ab')$_2$ anti-rat Fab. Both antibodies were conjugated with fluorescein (a gift of Dr S. V. Hunt). Lymph node cells were incubated with each of the two reagents and histograms of cell fluorescence obtained by flow cytofluorography using a Becton–Dickinson FACS II cell sorter. Using lymph node cells from a PVG *1-a* homozygous rat as a positive control, 37.1% of cells were shown to react with the anti-*1-a* allotype reagent and 39.4% with the anti-Fab antibody. For lymph node cells from a PVG donor the corresponding values were 1.8 and 35.1%, respectively. Leuk., leukaemic cells; BM, bone marrow cells.

* Died within 21 days of irradiation with no signs of leukaemia.

©1982Nature Publishing Group

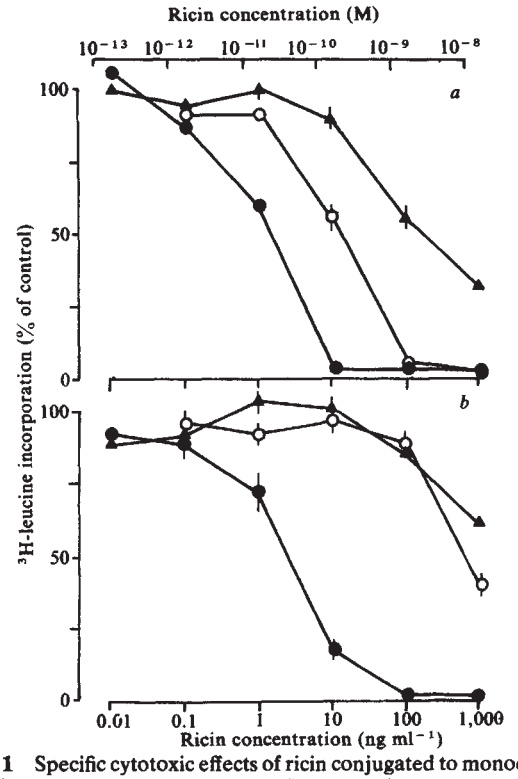


Fig. 1 Specific cytotoxic effects of ricin conjugated to monoclonal W3/25 antibody on PVG leukaemic cells in tissue culture. Blood was taken from a heavily leukaemic rat when its white blood cell count was 1.5×10^5 mm^{-3} and the leukaemic cells were separated by centrifuging the blood on Isopaque–Ficoll. The cells at a concentration of 2.5×10^6 ml^{-1} were incubated for 1 h at 37 °C with ricin (○), W3/25–ricin (●) or OX8–ricin (▲) in RPMI-1640 medium supplemented with 10% heat-inactivated fetal calf serum and antibiotics which either contained (*b*) or did not contain (*a*) 100 mM lactose. The cells were then washed in the same medium (but without toxin solutions) and incubated at 37 °C in fresh tissue culture medium without lactose. Their capacity to incorporate ^{3}H-leucine was measured 23 h later. Full details of the assay have been described previously[3]. The ^{3}H-leucine incorporation is expressed as a percentage of that in untreated control cultures which incorporated 10,300 c.p.m. per μCi. Each data point represents the geometric mean of three determinations, the standard deviations of which are indicated by vertical lines unless smaller than the points as plotted.

recipient and thus produce B lymphocytes that expressed the donor (*1-a*) allotype. The percentage chimaerism was compared with that obtained when untreated bone marrow cells, to which leukaemic cells had not been added, were injected into irradiated littermates of the recipients of the bone marrow/leukaemic cell mixtures. The evidence that the chimaerism assay measures stem cell activity is presented elsewhere[14]. The results of this experiment are presented in Table 1. All recipients of untreated bone marrow/leukaemic cell mixtures developed leukaemia before their level of B-cell chimaerism had been measured. None of the rats given W3/25–ricin-treated cell mixtures containing 10^3 leukaemic cells developed disease and of three recipients of inocula containing 10^4 leukaemic cells only one did so. All surviving rats developed a good level of chimaerism, although this was significantly lower in the recipients of bone marrow that had been exposed to the antibody–toxin conjugate than in control rats which had received bone marrow kept at 4 °C before injection. In a second experiment all recipients were lethally irradiated with 1,000 rad γ irradiation and survival of bone marrow stem cell activity was assayed, not by competition against surviving host bone marrow as in the first experiment but against a standard inoculum of 10^7 PVG bone marrow cells that were stored on ice until injection. As Table 1 shows, the results were very similar to those of the first experiment. Although acute radiation deaths limited the number of surviving animals in the group receiving 10^3 leukaemic cells, two out of three rats receiving 10^4 treated cells became long-term survivors with no sign of leukaemia and with significant chimaerism.

These experiments demonstrate that in the presence of 100 mM lactose the W3/25–ricin conjugate showed a high degree of target specificity. Approximately 99.9% of the leukaemic cells were killed by the conjugate while bone marrow stem cell activity was reduced by less than 50%. Some of the loss of stem cells can be attributed to effects other than ricin toxicity because incubation of the bone marrow in 100 mM lactose at 37 °C for 1 h in the absence of the antibody–ricin preparation resulted in a 13% reduction of chimaerism in assays similar to those described in Table 1.

The inclusion of lactose in the incubation medium not only prevented possible nonspecific killing of cells but had an important effect on the amount of conjugate bound nonspecifically to cells (such as erythrocytes) that was injected into the rats. Animals injected with 10^7 bone marrow cells that had been incubated for 1 h at 4 °C with the antibody–ricin conjugate at 2.1 μg ml^{-1} ricin in lactose-free medium and then washed twice died within 24 h of injection. Decreasing the concentration of conjugate to 0.21 μg ml^{-1} ricin prolonged the survival of recipient rats by only 24 h. However, when the cell incubations were carried out with conjugate in the presence of 100 mM lactose none of the recipient rats died or showed signs of ill health.

These results suggest that antibody–ricin conjugates might be used in conjunction with lactose to destroy malignant cells in human marrow transplants in situations where the antibody–A-chain conjugate is insufficiently effective.

Received 8 February; accepted 31 March 1982.

1. Olsnes, S., Saltvedt, E. & Pihl, A. *J. biol. Chem.* **249**, 803–810 (1974b).
2. Youle, R. J. & Neville, D. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5483–5486 (1980).
3. Thorpe, P. E. *et al. Clin. exp. Immun.* **43**, 195–200 (1981).
4. Dibley, M., Dorsch, S. & Roser, B. *Pathology* **7**, 219–235 (1975).
5. Mason, D. W. & Williams, A. F. *Biochem. J.* **187**, 1–20 (1980).
6. Williams, A. F., Galfre, G. & Milstein, C. *Cell* **12**, 663–673 (1977).
7. White, R. A. H., Mason, D. W., Williams, A. F., Galfre, G. & Milstein, C. *J. exp. Med.* **148**, 664–673 (1978).
8. Brideau, R. J., Carter, P. B., McMaster, W. R., Mason, D. W. & Williams, A. F. *Eur. J. Immun.* **10**, 609–615 (1980).
9. Barclay, A. N. *Immunology* **42**, 593–600 (1981).
10. Thorpe, P. E. & Ross, W. C. J. *Immun. Rev.* **62**, 119–158 (1982).
11. Krolick, K. A., Uhr, J. W. & Vitetta, E. S. *Nature* **295**, 604–605 (1982).
12. Thorpe, P. E., Edwards, D. C., Ross, W. C. J. & Davies, A. J. S. in *Monoclonal Antibodies in Clinical Medicine* (eds Fabre, J. & McMichael, A.) (Academic, New York, in the press).
13. Olsnes, S. & Pihl, A. in *Pharmacology of Bacterial Toxins* (eds Drew, J. & Dorner, F.) (Pergamon, Oxford, in the press).
14. Hunt, S. V. & Fowler, M. H. *Cell Tissue Kinetics* **14**, 445–464 (1981).

Specificity of UV mutagenesis in the *lac* promoter of M13*lac* hybrid phage DNA

J. Eugene LeClerc & Nancy L. Istock

Department of Biochemistry, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, USA

The mutagenic effect of UV irradiation on *Escherichia coli* and its phages depends on induction of the *E. coli recA*$^+$ *lexA*$^+$ regulatory system (for review, see ref. 1). It is not understood how mutations result from UV damage in DNA, particularly with respect to the function(s) of proteins that may be induced in UV-irradiated cells. Misincorporation of nucleotides during DNA synthesis on template DNA containing UV-induced lesions may occur predominantly at the specific sites of damage, implying that the DNA replication apparatus in UV-irradiated cells can bypass premutational lesions. Alternatively, misincorporation may occur randomly, suggesting an altered and inaccurate DNA replication apparatus in UV-irradiated cells. To

0028–0836/82/240596—03$01.00

© 1982 Macmillan Journals Ltd

study the specificity of UV mutagenesis at the nucleotide level, we have used M13*lac* hybrid phage, developed by Messing and colleagues[2,3] for rapid nucleotide sequencing of cloned DNA, as a forward mutation system. Nucleotide changes in a nonessential *lac* DNA sequence of the hybrid phage genome were determined by sequencing the DNA of mutant clones derived from UV-irradiated hybrid phage grown in UV-irradiated cells. We report here the nucleotide changes observed in the *lac* promoter of the mutant hybrid phage.

M13*lac* hybrid phage contains the M13 single-stranded DNA genome with an insert of the regulatory region and part (α-peptide) of the *lacZ* gene for lactose metabolism in *E. coli*. Infection of *E. coli* JM103 *lacZΔM15* cells with M13*lac* hybrid phage allows production of a functional β-galactosidase by intracistronic complementation. Infected cells are therefore lac^+, detected as blue plaques on indicator plates[2]. Inactivation of α-complementing activity, for example, by insertion of foreign DNA into the $lacZ'(\alpha)$ gene of the hybrid phage genome, gives rise to colourless plaques of infected cells, easily distinguishable from the blue plaques of M13lac^+-infected cells. In the studies reported here, inactivation of α-complementing activity was observed as a consequence of mutagenesis of UV-irradiated M13*lac* hybrid phage grown in UV-irradiated cells. Forward mutations in M13*lac* DNA were detected in the nonessential *lac* gene of the hybrid phage, so that all DNA changes (nucleotide substitutions, additions or deletions) resulting from UV mutagenesis could be observed by sequencing mutant clones.

Table 1 shows the mutation frequencies for induction of mutant *lac* hybrid phage derived from M13mp2 (ref. 3). Irradiation of M13mp2 hybrid phage at a UV dose giving ~0.01% survival, or irradiation of *E. coli* JM103 host cells at a UV dose giving maximal UV reactivation of phage (N.L.I. and J.E.L., unpublished results) showed no statistically significant enhancement of mutation frequency over that for spontaneous production of mutant *lac* hybrid phage. UV irradiation of both phage and host cells caused an ~12-fold increase in mutation frequency. Direct plating of irradiated phage ensured that mutant clones were of independent origin. Mutant plaque phenotypes ranged from colourless to a medium blue that was detectably different from the blue plaques of M13mp2-infected cells on indicator plates, thus the colour reaction allowed detection of *lac* mutants completely or partially defective in α-complementing activity; 115 mutant clones resulting from UV irradiation of both phage and host cells were picked and purified for analysis of nucleotide changes.

Table 1 Induction of *lac* mutations in M13mp2 hybrid phage

UV dose (J m^{-2}) Phage	Cells	Mutation frequency
0	0	$3.1\ (\pm1.6)\times10^{-4}$
125	0	$7.1\ (\pm1.3)\times10^{-4}$
0	50	$3.6\ (\pm5.1)\times10^{-4}$
125	50	$3.9\ (\pm0.2)\times10^{-3}$

E. coli JM103 were grown to 2×10^8 cells ml^{-1} at 37 °C in YT medium[13], then washed and resuspended in buffer (*p*H 8.0) consisting of 10 mM Tris, 67 mM KCl, 17 mM NaCl, 1 mM $MgSO_4$ and 1 mM $CaCl_2$. Cells were UV-irradiated at a dose rate of 0.8 J m^{-2} s^{-1} with stirring on ice. Cells were concentrated fivefold and resuspended in YT medium. Phage were diluted to 10^8 ml^{-1} in buffer and UV irradiation was carried out as for the cells. Cells were infected at multiplicity of ≤0.005 for 5 min at room temperature, plated onto indicator plates[2] and incubated overnight at 37 °C. All operations were carried out under yellow light. Mutation frequencies (±s.d.) are given as the ratio of plaques having mutant phenotypes to total plaques.

The single-stranded DNA of mutant hybrid phages was sequenced using the chain termination method[4]; 40 of the 115 mutant DNAs sequenced showed nucleotide changes in the *lac* promoter of M13mp2 DNA. In Fig. 1, autoradiograms of nucleotide sequencing gels show examples of base substitutions and deletions identified in the *lac* promoter. Figure 2 shows a compilation of all sequence changes (viral strand) detected. Most of these mutations are clustered in three regions of the M13mp2 *lac* promoter: the −10 region or Pribnow box, the −35 region, and the binding site for the CAP protein. Single base substitutions account for over 80% of the changes induced by UV mutagenesis. Tandem base changes occurred at two sites and non-tandem base changes were found in one mutant clone (see Fig. 2 legend). Single base deletions occurred at three sites, although assignment of the exact positions of the deletions was impossible because they occurred at TT or CCCC sites. No nucleotide additions have yet been observed in UV-induced mutants.

Mutagenesis of the *lac* portion of M13mp2 hybrid phage is most enhanced by UV irradiation both of the host cells, to induce the cellular mutagenesis system, and of the phage, containing the target DNA molecule (Table 1). As pyrimidine dimers are the major and potentially lethal photoproducts induced by UV irradiation of DNA[5,6] the nucleotide changes resulting from UV mutagenesis of M13*lac* hybrid phage were analysed with respect to sites of potential pyrimidine dimer formation. Of all changes observed, 77% occurred at such sites, although most mutations were found in the −35 region where promoter mutations are expected (ref. 7 and references therein) and which is pyrimidine-rich. These sites do represent hotspots for UV mutagenesis, however, because preliminary analysis of collections of *lac* mutants induced by chemical mutagens has revealed additional nucleotide changes at positions −31, −45, −55 and −56 that were not observed as a consequence of UV irradiation (data not shown). All three nucleotide deletions occurred at adjacent pyrimidines. For base substitution

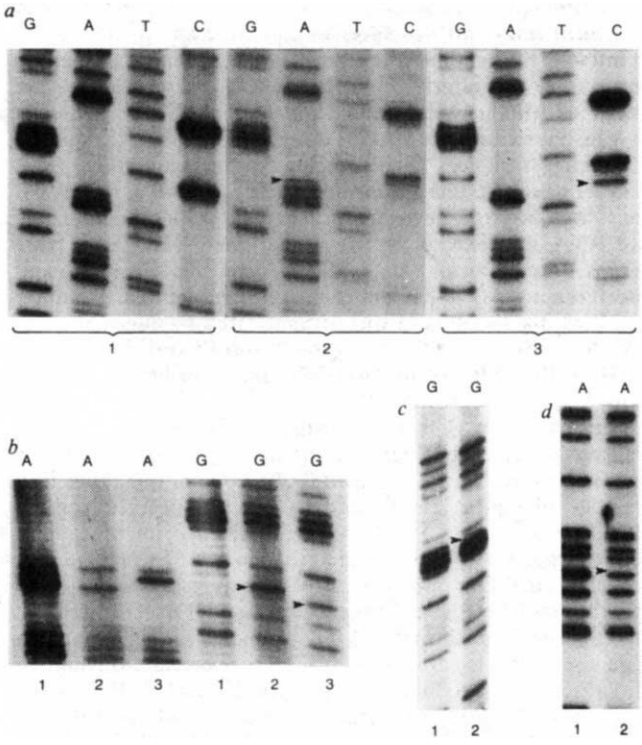


Fig. 1 Autoradiograms of nucleotide sequencing gels using templates prepared from M13mp2 hybrid phage and *lac* mutant derivatives. Preparation of templates, sequencing reactions and gel electrophoresis were carried out as described elsewhere[14] except that a 15-nucleotide primer (P.L. Biochemicals) was used. In each panel, lane 1 shows the sequence of normal M13mp2 DNA for comparison. The positions of nucleotide substitutions or deletions in mutant DNAs are indicated by arrowheads. *a*, C→T (viral strand) transition at position −37 (2) and T→G (viral strand) transversion at position −36 (3). *b*, T→C (viral strand) transitions at positions −35 (2) and −34 (3). *c*, C (viral strand) deletion (see text). *d*, T (viral strand) deletion (see text).

© 1982 Nature Publishing Group

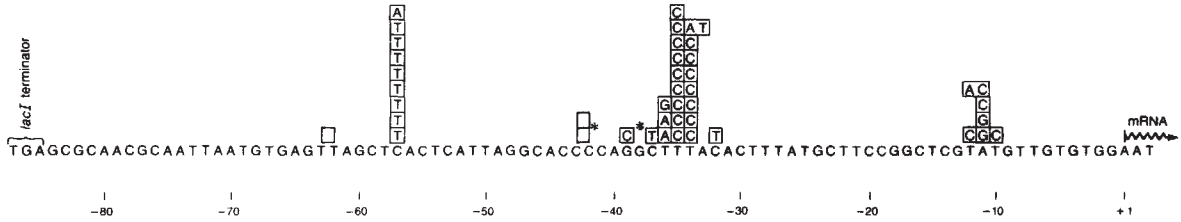


Fig. 2 Nucleotide changes in the *lac* promoter of mutant M13mp2 hybrid phage. The *lac* promoter sequence in hybrid phage DNA is numbered from the RNA polymerase initiation site. Each box represents a nucleotide change detected in *lac* mutant clones; nucleotide substitutions are shown in the box and nucleotide deletions are represented by open boxes. Tandem nucleotide substitutions are shown at positions −11, −12, and −34, −35. The exact positions of nucleotide deletions cannot be determined (see text). * Non-tandem nucleotide changes found in one mutant clone.

mutations at potential pyrimidine dimer sites, transitions were favoured over transversions (5 : 1), in agreement with previous studies on the specificity of UV mutagenesis in *E. coli*[8] and yeast[9]. It is particularly interesting that no tandem base substitutions occurred at potential pyrimidine dimer sites; the small proportion observed were at TA sites. Random substitution of nucleotides at pyrimidine dimer sites should yield up to 56% of changes at adjacent pyrimidines as tandem double changes, although the bias for transitions in UV mutagenesis should decrease that frequency.

Brandenburger *et al.*[10] determined the DNA sequences of revertants of M13 amber mutant phages grown in conditions of UV irradiation similar to those reported here. They concluded that pyrimidine dimers are not the main sites of base substitution mutagenesis in UV-irradiated M13 DNA. Unlike the M13*lac* forward mutation system, the greatest enhancement of mutations in the M13 reversion system is observed by UV-irradiating the host cells alone, so that nucleotide changes would not be expected to occur at potential pyrimidine dimer sites. In the case of UV irradiation of M13*am7H3* grown in UV-irradiated bacteria, however, Brandenburger *et al.*[10] did observe a fourfold increase in reversion frequency over that induced by UV irradiation of host cells alone, and the sequencing data showed a spectrum of mutations distinctly shifted to nucleotide changes at the 3′ pyrimidine of adjacent thymines.

Clearly, mutagenesis in UV-irradiated cells occurs at sites not directly associated with pyrimidine dimers. The principal sites of misincorporation may be at pyrimidine dimers, however, in which case these lesions in DNA templates are not totally non-instructive (compare ref. 8). The paucity of UV-induced tandem base changes at adjacent pyrimidines in mutant M13*lac* hybrid phage, M13 phage[10], *E. coli*[11] and yeast[12] supports this view.

We thank Diana Bongartz for technical assistance and Chris Lawrence for helpful discussions. This work was supported by a Cottrell Research Grant from Research Corporation, an ACS Institutional Research Grant, and NIH grant GM 27817.

Received 4 January; accepted 5 April 1982.

1. Witkin, E. M. *Bact. Rev.* **40**, 869–907 (1976).
2. Messing, J., Gronenborn, B., Müller-Hill, B. & Hofschneider, P. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3642–3646 (1977).
3. Gronenborn, B. & Messing, J. *Nature* **272**, 375–377 (1978).
4. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
5. Setlow, J. K. & Setlow, R. B. *Nature* **197**, 560–562 (1963).
6. Setlow, R. B. & Setlow, J. K. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1250–1257 (1962).
7. Siebenlist, U., Simpson, R. B. & Gilbert, W. *Cell* **20**, 269–281 (1980).
8. Lawrence, C. W. *Molec. gen. Genet.* **182**, 511–513 (1981).
9. Lawrence, C. W. & Christensen, R. B. *Molec. gen. Genet.* **177**, 31–38 (1979).
10. Brandenburger, A. *et al. Nature* **294**, 180–182 (1981).
11. Coulondre, C. & Miller, J. H. *J. molec. Biol.* **117**, 577–606 (1977).
12. Lawrence, C. W., Stewart, J. W., Sherman, F. & Christensen, R. *J. molec. Biol.* **85**, 137–162 (1974).
13. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
14. Heidecker, G., Messing, J. & Gronenborn, B. *Gene* **10**, 69–73 (1980).

Expression of human β-interferon cDNA under the control of a thymidine kinase promoter from herpes simplex virus

Gregory R. Reyes*, Elizabeth R. Gavis*, Alexander Buchan*, N. Babu K. Raj†, Gary S. Hayward* & Paula M. Pitha†

* Department of Pharmacology and Experimental Therapeutics and † Oncology Center, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

The availability of an isolated interferon gene cDNA clone and biological assays for its expression, provide an opportunity to examine the functional role of sequences preceding the coding portion of the interferon gene. In the present study, we examined whether the human β-interferon (IFN-β) gene can be transcribed and induced efficiently after replacing the normal 5′ flanking sequences[1,2] with a defined promoter segment from another gene. For this purpose the coding portion of the IFN-β gene[3] was inserted downstream from the promoter sequence and RNA start site of the thymidine kinase (TK) gene from herpes simplex virus type 1 (HSV-1; ref. 4). Expression of interferon activity under the control of the viral TK promoter in this chimaeric plasmid was demonstrated by microinjection into the nuclei of *Xenopus* oocytes[5,6] and by transfection into mouse cells[7] followed by superinfection with herpes simplex virus.

The interferon gene is particularly suitable for functional analyses because the high specific activity of the interferon protein product permits detection of picogram quantities in a sensitive biological assay[8], and its species specificity allows identification of the interferon synthesized from the added heterologous DNA as opposed to that encoded by endogenous host cell genes. Our objectives were both to insert the 560-base pair (bp) structural gene sequence of the pHFβ-cDNA clone[9] downstream from the HSV-1 TK promoter sequence, and to have a second selectable intact TK marker present to facilitate the subsequent introduction of the hybrid TK–IFN cDNA gene into Ltk⁻ cells. The construction of plasmids containing the TK promoter fused to the coding sequences of interferon cDNA in both the sense (pGR192) and nonsense (pGR191) directions is summarized in Fig. 1, together with the principal features of the parent pGR18, pGR156 and pHFβ-cDNA plasmids. In pGR192, the coding (or sense) strand of the interferon insert lies downstream from the HSV-1 TK promoter, 54 nucleotides after the 5′ transcription initiation site but in front of the first AUG for the TK protein[4,10].

To determine whether the HSV TK flanking sequences can promote valid transcription of the IFN-β gene, plasmids containing the hybrid TK–IFN cDNA gene were tested for the ability to direct synthesis of interferon after microinjection into

0028–0836/82/240598—04$01.00
© 1982 Macmillan Journals Ltd

nuclei of *Xenopus* oocytes. We compared the levels of interferon synthesized in oocytes injected with DNA from the pHFβ-cDNA, pGR191 and pGR192 clones with those from oocytes injected with pBR322 and pGR18 (Table 1). In both experiments, oocytes receiving the pGR192 clone produced interferon (120–530 U ml^{-1}), whereas no interferon activity (< 20 U ml^{-1}) could be detected with any of the other plasmids. The levels of interferon obtained were comparable to those synthesized in oocytes after microinjection with poly(A)-containing RNA isolated from human fibroblasts after induction with poly(rI·rC)[11,12]. These results indicate that efficient synthesis of interferon mRNA occurred only when the cDNA coding sequence was inserted in the correct orientation downstream from the TK promoter.

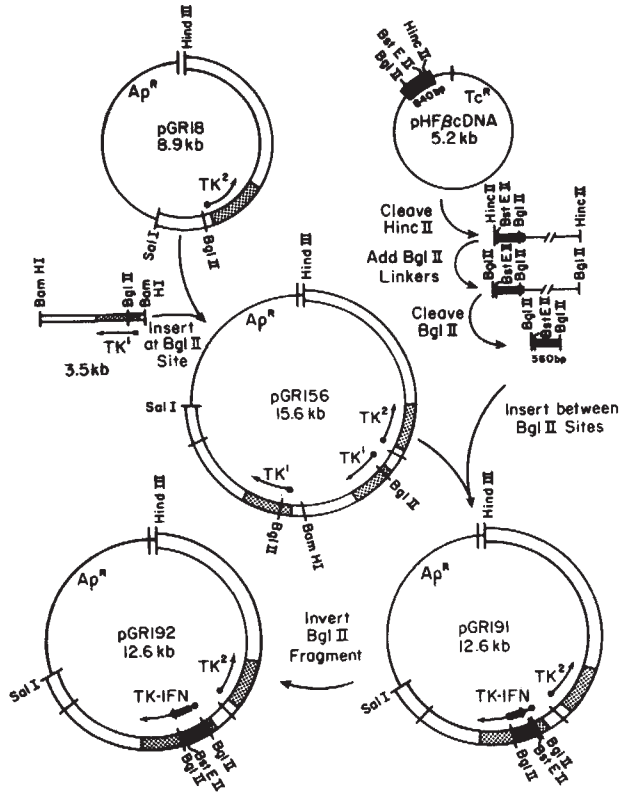


Fig. 1 Diagram illustrating the general strategy used to construct an interferon expression plasmid containing the coding portion from cloned IFN-β cDNA linked to the TK gene. The pGR18 plasmid[26] contains a cloned 4.8-kb *Hin*dIII–*Sal*I fragment from HSV-2(333) viral DNA (open bars), which includes the HSV-2 TK gene (stippled bars), inserted into pBR322 (single line). The 3.5-kb *Bam*H1–O fragment shown contains the HSV-1 TK gene and was derived originally from the HSV-1(MP) strain[6]. Arrows indicate the direction of transcription in each of the TK genes and solid circles represent the promoter regions. The pHFβ-cDNA clone described previously[9] was derived from a cDNA library of human fibroblast cells induced with poly(rI·rC) and was inserted into the *Pst*I site of pBR322. The 840-bp human fibroblast cDNA segment of this plasmid (solid bars) possesses single *Bgl*II, *Bst*EII, *Hinc*II, *Pst*I, *Pvu*II and *Taq*I cleavage sites in the same arrangement as those reported by Derynck *et al.*[3] for their 860-bp IFN-β cDNA plasmid. We have used only the 560-bp *Hinc*III–*Bgl*II fragment encoding the IFN-β structural gene which includes the AUG translation initiation codon for the signal peptide and the UGA termination signal. To appropriately modify the 560-bp interferon fragment phosphorylated *Bgl*II linkers[27] were ligated at 20-fold molar excess to the *Hinc*II-cleaved pHFβ-cDNA plasmid at 10 °C for 30 h in a reaction volume of 50 μl (refs 28, 29). Gel electrophoresis of a series of time point samples confirmed that linkers were ligated to the *Hinc*II fragments, and that formation of higher multimers of the *Hinc*II fragments with and without linker attachment had also occurred. After ligation, the DNA was ethanol-precipitated and resuspended in 0.01 M Tris-HCl, 0.001 M EDTA *p*H 8.0. Appropriate amounts of $MgCl_2$ and NaCl were added in order for *Bgl*II digestion to generate cohesive extensions at the 5′ end of the interferon cDNA and also to cleave at the *Bgl*II site at the 3′ end of the coding sequence. This fragment was then inserted between the two *Bgl*II sites in the two tandemly repeated HSV-1 TK genes within pGR156. Cultures of *Escherichia coli* HB101 were transformed by the calcium shock method of Mandel and Higa[30]. For hybridization screening, the transfection mixture was spread on to selective ampicillin plates and grown until the colonies reached a size of 0.1–0.2 mm. The colonies were then picked up on the surface of keyed Whatman No. 1 filter paper circles and transferred (colonies uppermost) to chloramphenicol plates for further incubation (12–48 h). The filters were treated with alkali, neutralized, washed and baked essentially as described by Hanahan and Meselson[31]. Positive clones on these filters were detected directly by hybridization with appropriate nick-translated probes. Digestion of selected plasmids containing a single 560-bp *Bgl*II insert with *Pst*I, *Pvu*II or *Bst*EII, followed by hybridization to Southern transfers from agarose gels, revealed that the interferon sequence in pGR191 was oriented in the opposite direction to that of the TK promoter, whereas that in pGR192 represented the correct orientation.

The presence of an intact HSV-2 TK gene in our expression plasmids was utilized to establish permanent TK$^+$ cell lines by calcium transfection of Ltk$^-$ cells and selection of colonies growing in hypoxanthine–aminopterin–thymidine (HAT) medium[13]. Five TK$^+$ clones from the pGR191 transfections and five from the pGR192 transfections were tested, but none produced interferon constitutively (detection limit 1 U ml^{-1}). As the expression of an active resident viral TK gene in mouse L-cell lines has been shown to be enhanced 5–10-fold by superinfection with HSV-1 (*tk*$^-$) virus[14], we reasoned that the HSV-1 TK promoter attached to the interferon cDNA, although apparently inactive for constitutive synthesis in our cell lines, might nevertheless respond to the viral regulatory functions introduced by a superinfecting viral genome. Table 2 shows that the LH$_2$p192–8 subclone does indeed express interferon activity induced by HSV-1 and HSV-2, but not by infection with more distantly related viruses of the herpes group, for example, cytomegalovirus (CMV) or pseudorabies virus (PRV).

The interferon synthesized both in microinjected oocytes and LH$_2$p192–8 cells displayed the same species specificity and antigenicity as that derived from human fibroblasts, that is, it was active on human cells but not on bovine or mouse cells, and it was completely neutralized by anti-human IFN-β antiserum[15], but not by anti-human α-interferon[16] or anti-mouse IFN-α plus -β antisera[17]. We conclude that the interferon synthesized is authentic human IFN-β, which could only have been encoded by the exogenous human IFN-β sequences introduced into these mouse cells by transfection.

The kinetics of IFN-β synthesis in superinfected LH$_2$p192–8 cells, including initiation within 2–3 h, maximal expression at 6 h and a decrease by 20 h, follows the expected time course for viral transcriptional regulation of a delayed-early promoter, that is, induction by immediate-early functions followed by repression after synthesis of late functions. Additional experiments with a viral *ts* mutant that synthesizes an inactive immediate-early protein supported this conclusion (Table 2, expt 2). The *ts*B2 mutation maps in the HSV-1 IE gene IV, which encodes a phosphorylated protein of molecular weight 175,000 (175K) having DNA binding properties and a nuclear chromatin-associated location[18,19]. This protein has been shown

Table 1 Human IFN-β synthesized in *Xenopus* oocytes

Plasmid injected	DNA (μg μl^{-1})		Interferon (U ml^{-1}) GM2504		Interferon (U ml^{-1}) Ly	
	a	*b*	*a*	*b*	*a*	*b*
pHFβ-DNA	0.20,	0.10	<20,	<20	<20,	—
pGR192	0.10,	0.10	530,	120	<20,	—
pGR191	0.17,	0.10	<20,	<20	<20,	—
pBR322	0.14,	0.10	<20,	<20	<20,	—
pGR18	0.15,	—	<20,	—	<20,	—

Mature *Xenopus* oocytes were prepared as described by McKnight and Gavis[6] and a sample (3–5 ng) of supercoiled DNA in a 30-nl volume was microinjected into each nucleus. Fifty oocytes per assay were washed and incubated in 3 ml of Barth's medium at 18 °C for 36 h. Interferon activity in the medium was assayed as described previously[11,12] using the cytopathic assay of Finter[24] on either human fibroblast cells trisomic for chromosome 21 (GM2504) or on mouse Ly cells. The interferon titres are given in international reference standard units (69/17). One unit of standard human IFN-β titrated as 1 unit in our assay on GM2504 cells and as 10^{-3} units on Ly cells. The results of two independent experiments are shown: *a*, *Xenopus borealis* oocytes; *b*, *Xenopus laevis* oocytes. —, Not determined.

© 1982 Nature Publishing Group

to have a direct role, during HSV infection, in the transcriptional activation (or depression) of viral delayed-early genes including the thymidine kinase gene[20,21]. Infection of LH_2p192–8 cells with HSV-1 (*ts*B2) in non-permissive conditions (39.5 °C) resulted in a 50-fold reduction in the level of IFN-β synthesis compared with infection by the same virus in permissive conditions (33.5 °C).

Induction of interferon activity was not detected after HSV-1 infection in five LH_2p191 cell lines and in four other LH_2p192 cell lines when tested at sixth passage (only one of each is shown in Table 2). Surprisingly, in earlier experiments with these cell lines at third passage, we found that human IFN-β could be induced in eight of the ten LH_2p191 or LH_2p192 clones by either poly(rI·rC) treatment or Newcastle disease virus infection, indicating that the IFN-β cDNA sequences were present and that they also respond to induction by some other mechanism that does not require the presence of the TK promoter[13]. The inability to induce interferon in some LH_2p192 lines after HSV infection could represent either sequence rearrangements or loss of the TK–IFN hybrid gene after integration. To examine this question, DNA samples from LH_2p191 and LH_2p192 cell lines were digested with *Pvu*II (which cleaves once within the IFN-β structural sequence) and hybridized with a ^{32}P-labelled fragment probe consisting of the isolated 560-bp IFN-β coding sequence (Fig. 2). The single-copy IFN-β gene sequences present in human fibroblast cell DNA gave two bands of the size expected from previous reports[22,23]. Similarly, the LH_2p192–8 and LS_2p191-1 cell DNAs each exhibited two bands that correlated with the appropriate fragment sizes expected after cleavage at the flanking *Pvu*II sites in the HSV-2 TK gene. In contrast, DNA from the LH_2p192-1 line at sixth passage did not hybridize with the interferon cDNA probe. Therefore, the lack of inducibility observed in the LH_2p192-1 line correlated with the absence of IFN-β sequences from the cellular DNA, whereas the lack of induction in the LH_2p191-3 line was probably related to the reverse orientation of the promoter with respect to the IFN-β coding sequence. These data strongly support the notion that the synthesis of IFN-β in the LH_2p192-8 transfected mouse cell line represents transcriptional events related directly to the correct positioning of the TK promoter sequence.

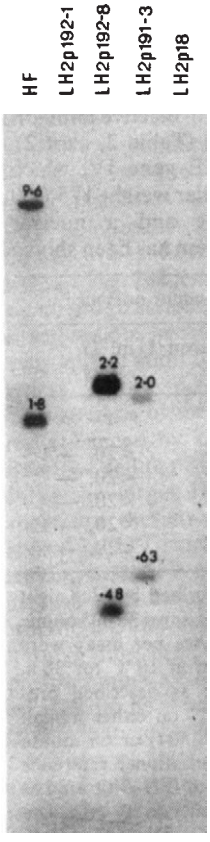


Fig. 2 Demonstration of the presence of human IFN-β DNA in transfected mouse cells. High molecular weight DNA was extracted from mass cultures of appropriate subcloned cell lines at sixth passage, digested to completion with *Pvu*II and fractionated by agarose gel electrophoresis (1.2% horizontal gels, 10 μg DNA per sample). The DNA was transferred to nitrocellulose and incubated with a ^{32}P-labelled probe prepared by nick-translation of the isolated 560-bp IFN-β fragment from pGR191. Hybridization was carried out at 37 °C for 18 h in 50% formamide, 10% dextran sulphate, 1×Denhardt's mix, 0.5% SDS, 100 μg ml^{-1} denatured salmon sperm DNA, 20 μg ml^{-1} *E. coli* RNA, 3×SSC, 1 mM EDTA, and 50 mM Tris-HCl *p*H 7.6, by a procedure based on that described elsewhere[32]. The sizes of hybrid bands detected in the autoradiograph are given in kilobase pairs. Lane 1, human fibroblast HF-4 cell DNA used as a single-copy reference; lane 2, LH_2p192-1 DNA; 3, LH_2p192-8 DNA; 4, LH_2p191-3 DNA; 5, LH_2p18 DNA.

Table 2 Induction of human IFN-β in transfected mouse cells

Cell line	Virus used	Interferon induced (U ml^{-1}) Not infected*	Superinfection† 3 h	6 h	20 h
Expt 1					
Ltk^-	HSV-1(MP), 36 °C	<5	<5	<5	<5
LH_2p18	HSV-1(MP), 36 °C	<5	<5	<5	<5
LH_2p191-3	HSV-1(MP), 36 °C	<5	<5	<5	<5
LH_2p192-1	HSV-1(MP), 36 °C	<5	<5	<5	<5
LH_2p192-8	HSV-1(MP), 36 °C	<5	60	160	45
Expt 2					
Ltk^-	HSV-1(MP), 36 °C	<5	—	<5	—
LH_2p192-8	HSV-1(MP), 36 °C	<5	—	320	80
LH_2p192-8	HSV-1(KOS*ts*B_2), 33.5 °C	—	—	720	40
LH_2p192-8	HSV-1(KOS*ts*B_2), 39.5 °C	—	—	15	<5
LH_2p192-8	HSV-2(333), 36 °C	—	—	720	—
LH_2p192-8	CMV(Towne), 36 °C	—	—	<5	<5
LH_2p192-8	PRV, 36 °C	—	—	<5	—
HF-4	HSV-1(MP), 36 °C	<5	—	<5	—

The transfection protocol used was essentially that described by Graham *et al.*[25] as modified for TK selection by Wigler *et al.*[7] and is described in detail elsewhere[13]. Ltk^- cells are thymidine kinase-negative L cells; LH_2p18, LH_2p191 and LH_2p192 cells are TK^+ clones selected in HAT medium after transfection of Ltk^- cells with the pGR18, pGR191 and pGR192 plasmids, respectively. HF-4 cells are diploid human fibroblasts.

* Media from 5×10^5 cells per 35 mm dish were examined directly for the constitutive production of interferon (24 h collection) on human GM2504 indicator cells.

† Cell cultures were superinfected with wild-type herpesviruses at 36 °C or with HSV-1(KOS*ts*B_2) at either 33.5 °C (permissive conditions) or 39.5 °C (non-permissive conditions). The multiplicities of infection ranged from 10 plaque forming units per cell in expt 1 to 30 plaque forming units per cell in expt 2. The levels of interferon accumulated in the medium at various times after virus absorption were assayed on GM2504 cells.

By placing the interferon gene behind the TK promoter, we hoped that an active product might be expressed continuously in transfected cells in the same way that the TK gene product is expressed continuously in TK^+ cell lines. However, this was not the case. Although we selected for constitutive expression of the viral thymidine kinase enzyme behind one TK promoter in the LH_2p192-8 cell line, a second TK promoter introduced into the same cells, but controlling the unselected interferon cDNA gene, was apparently not expressed to any significant degree. Nevertheless, this second TK promoter was still available for transcriptional activation when the appropriate viral factors entered the cell. Therefore, the control of expression from the resident unselected TK promoter in the transfected mouse cells follows the normal pattern seen in HSV-infected cells. The fact that HSV infection does not induce interferon in a human fibroblast cell line (HF-4), and the lack of induction by the HSV-1 *ts*B2 mutant in LH_2p192-8 cells at non-permissive temperature, provided further evidence that it was neither the input virions nor viral DNA which activated the TK–IFN hybrid gene, but rather a specific HSV regulatory gene product (probably the 175K protein itself).

We thank Dr W. Fiers for the gift of the HF12cDNA clone used in the original identification of our pHFβ-cDNA plasmid, Drs K. Berg and E. DeMaeyer and J. DeMaeyer for gifts of antisera, and Dr B. Sollner-Webb for making available microinjection facilities. G.R.R. was a predoctoral fellow in the Medical Scientist Training Program at the Johns Hopkins School of Medicine (GM 07309). A.B. was supported by a NATO exchange program fellowship. This work was funded by NIH grants 5ROI AI-10944-09 (P.M.P.) and 2ROI CA-22130-04 (G.S.H.).

Received 5 January; accepted 26 March 1982.

1. Nagata, S., Mantei, N. & Weissman, C. *Nature* **287**, 401–408 (1980).
2. DeGrave, W., Derynck, R., Tavernier, J., Haegeman, G. & Fiers, W. *Gene* **14**, 137–143 (1981).
3. Derynck, R. *et al. Nature* **285**, 542–547 (1980).

© 1982 Nature Publishing Group

4. McKnight, S. L. *Nucleic Acids Res.* **8**, 5949–5964 (1980).
5. Mertz, J. E. & Gurdon, J. B. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1502–1506 (1977).
6. McKnight, S. L. & Gavis, E. R. *Nucleic Acids Res.* **8**, 5931–5948 (1980).
7. Wigler, M. *et al. Cell* **11**, 223–232 (1977).
8. Knight, E. Jr *J. Interferon Res.* **2**, 1–12 (1980).
9. Raj, N. B. K. & Pitha, P. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7426–7430 (1981).
10. McKnight, S., Gavis, E. R., Kingsbury, R. & Axel, R. *Cell* **25**, 385–398 (1981).
11. Reynolds, F. H. Jr, Premkumar, E. & Pitha, P. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4881–4885 (1975).
12. Raj, N. B. K. & Pitha, P. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1483–1487 (1976).
13. Pitha, P. M. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
14. Leiden, J. M., Buttyan, R. & Spear, P. *J. Virol.* **20**, 413–424 (1976).
15. Neurath, A. R., Strick, N., Raj, N. B. K. & Pitha, P. M. *J. Interferon Res.* **2**, 51–57 (1982).
16. Berg, K. *Scand. J. Immun.* **114**, 77–86 (1977).
17. DeMaeyer-Guignard, J., Tovey, M. G., Gresser, I. & DeMaeyer, E. *Nature* **271**, 622–625 (1978).
18. Cabral, G. A., Courtney, R. J., Schaffer, P. A. & Marciano-Cabral, F. *J. Virol.* **33**, 1192–1198 (1980).
19. Hay, R. T. & Hay, J. *Virology* **104**, 230–234 (1980).
20. Preston, C. M. *J. Virol.* **29**, 275–284 (1979).
21. Watson, R. J. & Clements, J. B. *Nature* **285**, 329–330 (1980).
22. Tavernier, J., Derynck, R. & Fiers, W. *Nucleic Acids Res.* **9**, 461–471 (1981).
23. Houghton, M. *et al. Nucleic Acids Res.* **9**, 247–266 (1981).
24. Finter, N. B. *J. gen. Virol.* **5**, 419–427 (1969).
25. Graham, F. L., Veldhuisen, G. & Wilkie, N. M. *Nature* **245**, 265–266 (1973).
26. Hayward, G. S., Reyes, G. R., Gavis, E. R. & McKnight, S. L. in *Herpesvirus DNA. Developments in Molecular Virology* Vol. 1 (ed. Becker, Y.) 271–306 (Martinus-Nijhoff, The Hague, 1981).
27. Sakonju, S., Bogenhagen, D. F. & Brown, D. D. *Cell* **19**, 13–25 (1980).
28. Dugaiczyk, A., Boyer, H. W. & Goodman, H. M. *J. molec. Biol.* **95**, 171–184 (1975).
29. Lillehaug, J. R., Kleppe, R. K. & Kleppe, K. *Biochemistry* **15**, 1858–1865 (1976).
30. Mandel, M. & Higa, A. *J. molec. Biol.* **53**, 159–162 (1970).
31. Hanahan, D. & Meselson, M. *Gene* **10**, 63–67 (1980).
32. Sager, R., Anisowicz, A. & Howell, N. *Cell* **23**, 41–50 (1981).

Unique insertion site of Tn*7* in the *E. coli* chromosome

Conrad Lichtenstein* & Sydney Brenner

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Among transposable elements, transposon Tn*7* is unusual in that it inserts into a particular site of the *Escherichia coli* chromosome with high efficiency[1,2] and with a unique orientation[2]. Most other transposable elements in prokaryotes are capable of translocating from and to a variety of sites in bacteriophage and plasmid DNA as well as in the chromosome[3]. We have previously shown that the 'attachment site' for Tn*7* in *E. coli* is located between *oriC* and *glmS* at minute 82 (ref. 4) of the chromosome[2]. We now report the DNA sequences of the termini of Tn*7*, of the attachment site (before and after Tn*7* insertion) and of an alternative site of Tn*7* attachment. We conclude that the specificity of the attachment site is determined by specific DNA sequences and that secondary structures may be involved.

Tn*7* is a very large (14 kilobase (kb)) element which codes for resistance to the antibiotics trimethoprim, streptomycin and spectinomycin[1]. Its preference for the attachment site is so pronounced that, when the site is available, virtually all transposition events are directed there[2]. It is not yet known whether the high efficiency with which Tn*7* transposes into *Pseudomonas aeruginosa*[5], *Klebsiella pneumoniae* and *Agrobacterium tumefaciens* (M. van Montagu, personal communication) involves a similarly unique site, but the transposition of Tn*554* in *Staphylococcus aureus* is known to involve such a site[6].

We have obtained the DNA sequence of the attachment site by the use of the plasmid pCPL6[2], previously constructed from the plasmid pBR322[7] by replacing the small *EcoRI–Bam*HI restriction fragment with the much larger (8×10^6 molecular weight) corresponding fragment derived from the transducing phage[8] λ *asn*5, which was shown to contain the Tn*7* attachment site[2]. The site-specific properties are retained[2]. Figure 1 shows a restriction map of pCPL6 including the Tn*7* insertion-site. For the DNA sequence of the non-preferred site, we have used the plasmid Co1E1::Tn*7*. In both cases, DNA sequences of appropriate restriction fragments were subcloned into an M13 phage vector and sequenced using the chain termination method[9].

*Present address: Department of Microbiology and Immunology, University of Washington, Seattle, Washington 98195, USA.

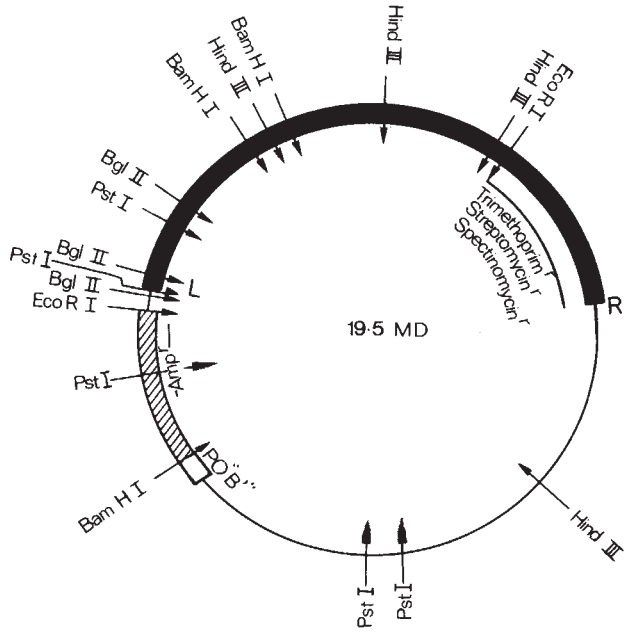


Fig. 1 Restriction map of pCPL6: :Tn*7*, showing the source of the sequenced material. A small *Bgl*II junction fragment at the left end (L) of Tn*7* was subcloned into the phage vector M13mp2/Bam (ref. 18). This fragment was used as a probe to identify, by plaque hybridization screening[19], the *Sau*3A restriction fragment from pCPL6 that includes the Tn*7* attachment site before insertion. The cloned insert of this latter phage was used similarly as a probe to identify phage re-combinants containing the *Sau*3A junction fragment of the right end (R) of Tn*7* in pCPL6::Tn*7*. This right-end junction fragment was then used as a probe to identify the *Sau*3a junction fragment at the right end of Tn*7* in the plasmid Co1E1::Tn*7*. In addition, a *Pst*I fragment overlapping with the left-end *Bgl*II junction fragment in pCPL6::Tn*7* was subcloned into the M13mp2/Pst phage vector[20]. The cloned inserts of all these M13 phage recombinants were sequenced by the chain terminator method[9] using the virion DNA as template[21] and a 17-bp synthetic primer (M. Gait, unpublished) or the 92-bp universal primer[21]. These DNA sequences are shown in Fig. 2. Note that the plasmid pCPL6::Tn*7* contains a secondary λ attachment site indicated as PO"B" in the figure (see ref. 2 for a complete description of the construction of pCPL6). —, *E. coli* DNA; ■, Tn*7* DNA; ▨, pBR322 DNA; □, λ DNA.

In two independent experiments, Tn*7* was inserted into the attachment site in pCPL6 by the techniques described in ref. 2; in each case, the site of attachment and the surrounding sequences were identical (see Fig. 2). The following properties of the DNA sequences are of interest.

As with many other transposons, Tn*7* insertions are bounded by short (5 base pair (bp)) directly repeated sequences. The occurrence of a duplication of the pre-existing 5-bp sequence must be a consequence of the mechanism of transposition.

The precise site of insertion occurs in both the *E. coli* chromosome and in Co1E1 within a short region of target DNA that contains inverted repeat sequences. Such sequences have alternative pairing configurations and may be represented either by linear duplex DNA or by looped-out cruciform or hairpin structures (see Fig. 3*a*, *b*).

It is difficult to evaluate the importance of these possible secondary structures in Tn*7* transposition. Thermodynamic calculations using the rules of Tinoco[10] suggest that in single-stranded DNA these structures can stably form at 25 °C ($\Delta G = -7.2$ kcal mol^{-1} for structure 3*a* and $\Delta G = -3.6$ kcal mol^{-1} for structure 3*b*). At 37 °C, however, the second of these free energy calculations is unlikely to be a sure guide to the stability of a cruciform structure. There is strong evidence that secondary structures exist in supercoiled duplex DNA (ref. 11), particularly in A+T-rich regions[12]. Bacteriophage λ, which shares many properties with transposable elements, has been shown to require supercoiled DNA for integration[13], and the bacterial chromosome is known to be composed of supercoiled domains[14]. Perhaps site selection by Tn*7* involves structures that directly result from this supercoiling. With other transposable elements, the site of insertion seems not to contain

0028-0836/82/240601—03$01.00

© 1982 Macmillan Journals Ltd

potential cruciforms, but there is a preference for insertion into A+T-rich regions[3].

Tn*7* has unusual asymmetrical termini. Each end contains regions that occur several times in the same orientation and which are closely homologous. They are contiguous in the left end of Tn*7* (boxes 2 and 3 of Fig. 2*a*). These direct repeated sequences of 22 bp were identified by computer analysis[23]; they are more than 75% homologous. In the right end of Tn*7*, they appear in the opposite orientation (boxes 2 and 3, Fig. 2*b*) but are separated by sequences not found at the other end. Figure 3*c* presents a potential hairpin structure that might form by bringing the homologous regions at each end of Tn*7* into register. The asymmetry at the ends of Tn*7* may account for its unique orientation of insertion at the attachment site, as other transposable elements have symmetrical termini and insert in either orientation[3].

Other transposable elements also possess directly repeated sequences within each end. For example, other members of the family of elements that generate 5-bp duplications have the sequence 5′ACGAAAA3′ followed by a poorly conserved direct repeat[3]. (A nearly homologous sequence, 5′ACAAAA3′, occurs several times within the ends of Tn*7* in box 3.) This common feature may have some role in recognition by the enzymes involved in transposition and, since the duplicated regions (that is, box 3) are more extensive in Tn*7*, this may in part account for the unusually efficient transposition of Tn*7*. Naturally, the nature of the target sequence must also determine the efficiency of transposition.

Tn*7* terminates in the dinucleotide sequence TG . . . CA, as has also been observed in bacteriophage Mu (ref. 3), eukaryotic transposable elements and the provirus forms of retroviruses[15].

A full understanding of the reasons for the extreme specificity of Tn*7* for its attachment site must await further studies. Clearly, the mere presence of a hairpin structure is insufficient, because many chromosomal and plasmid inverted repeat structures (some thermodynamically more stable than the preferred site) are not high efficiency target sequences. Tn*7* may show a preference for insertion near to regions showing DNA sequence homology with its termini, as both insertion sites reported here show ~65% homology with the terminal 12 bp at the ends of Tn*7*; such a preference has also been observed for other transposable elements[3]. Enzymes or mechanisms similar to those involved in bacteriophage integration may be involved, for the regions surrounding the Tn*7* attachment site and a region within one end of Tn*7* show homology to the core sequence of λ *att* (see Fig. 2). Such sequences also occur within the termini of other transposable elements[16], while the element IS*4* favours insertion into a site close to a secondary λ *att* in the *E. coli* chromosome[17].

A better understanding of the factors involved will require a mutational analysis of Tn*7* and its attachment site. Work is under way to determine the DNA sequences of mutations that affect this phenomenon and establish which nucleotides participate in site-specific Tn*7* integration.

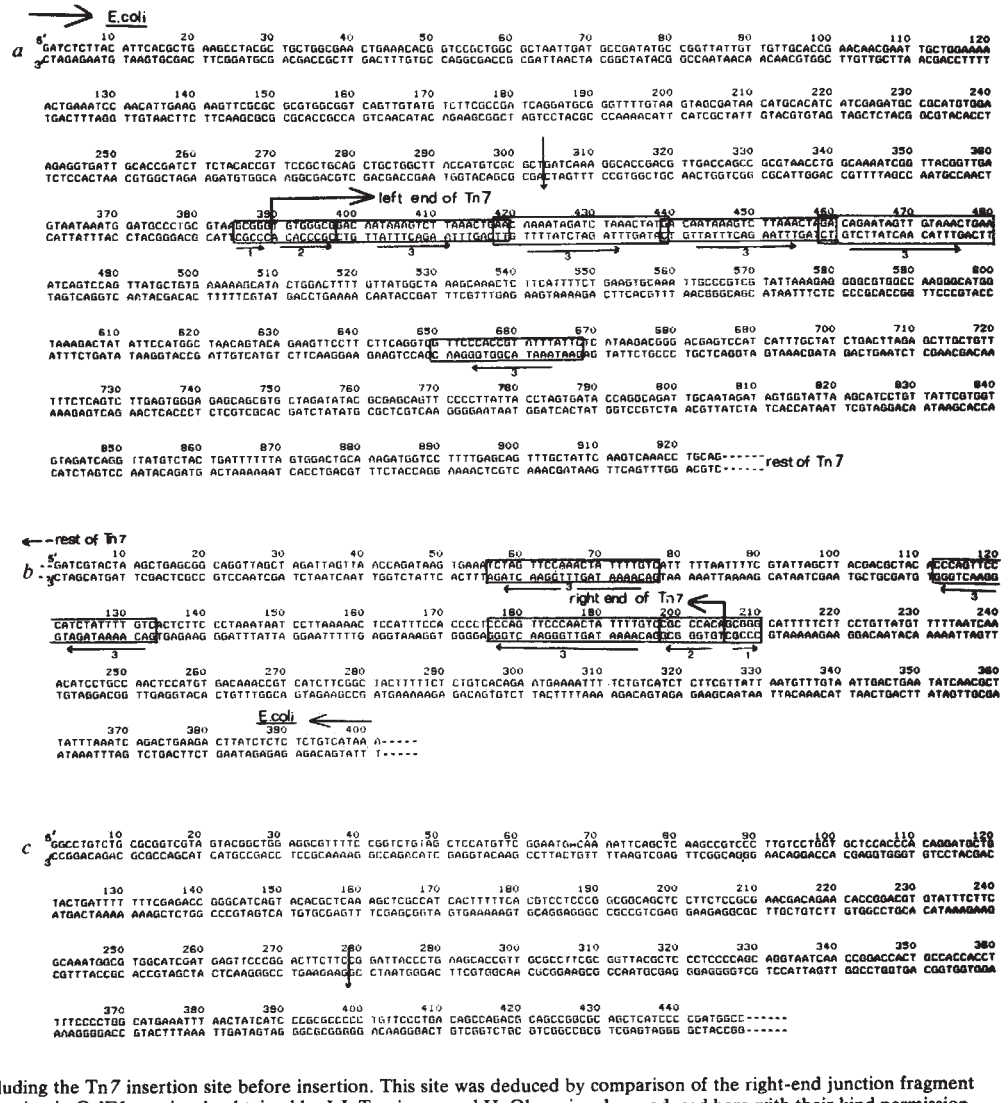


Fig. 2 *a*, *b*, The DNA sequence of the left and right ends of Tn*7* at its attachment site in the plasmid pCPL6::Tn*7*. It has been assumed that the insertion site of Tn*7* in the *E. coli* DNA region of this plasmid is identical to that of the Tn*7* attachment site in the context of the *E. coli* chromosome, an assumption substantiated by restriction enzyme analysis[2]. The left end of Tn*7* terminates at position 390 in *a*. The right end terminates at position 205 in *b*. The DNA sequence of a 13-kb region of Tn*7* DNA between these ends was not determined. Each end is flanked by *E. coli* DNA that surrounds the Tn*7* attachment site (positions 1–389 in *a* and 206–401 in *b*). The *Sau*3A restriction fragment (1 kb) from pCPL6, that includes the Tn*7* attachment site before insertion, begins at nucleotide 304 in *a* (as indicated by the arrow); this fragment was here shown to retain efficient site-specific properties with regard to Tn*7* transposition[2]. A 5-bp sequence of duplicated target DNA (indicated as box 1 in *a* and *b*) flanks the ends of Tn*7* as a direct repeat, as indicated by the arrows. A short G+C-rich region is present at each end of Tn*7* (box 2) in opposite orientation. Other regions of near direct repeat are indicated as box 3 at each end of Tn*7*. An asterisk at position 659 indicates an extra G:C base pair in one of these boxed regions. Certain regions show some DNA sequence homology to the core region (5′GCTTTTTTATACTAA3′) of λ *att* (ref. 16) when allowance is made for gaps at some nucleotide positions. These regions are at positions 210–225 (10/15 bp homology) 229–241 (11/15 bp homology) in *b*. A similar sequence occurs at positions 151–138 in the right end of Tn*7* in the opposite orientation (9/15 bp homology). Significant A+T-rich regions that surround the Tn*7* attachment site are at positions 343–369 in *a* (19/27 A+T) and 212–241 in *b* (24/30 A+T). A 12-bp DNA sequence at the Tn*7* attachment site (5′CGTAAGCGGGCA3′) shows an 8/12 homology with the terminal 12 bp of Tn*7*. *c* Shows the ColE1 DNA sequence of the *Hae*III G fragment[22], including the Tn*7* insertion site before insertion. This site was deduced by comparison of the right-end junction fragment of ColE1::Tn*7* with the sequence of this region in ColE1 previously obtained by J-I. Tomizawa and H. Ohmori and reproduced here with their kind permission. An arrow between positions 278–279 indicates the Tn*7* insertion site, which is in the same orientation as that shown in the attachment site. The sequence of the right end of Tn*7* in ColE1 was identical with that shown in *b*. Nucleotide positions 262–273 (*c*) show 7/12 bp homology with the terminal 12 bp of Tn*7*.

© 1982 Nature Publishing Group

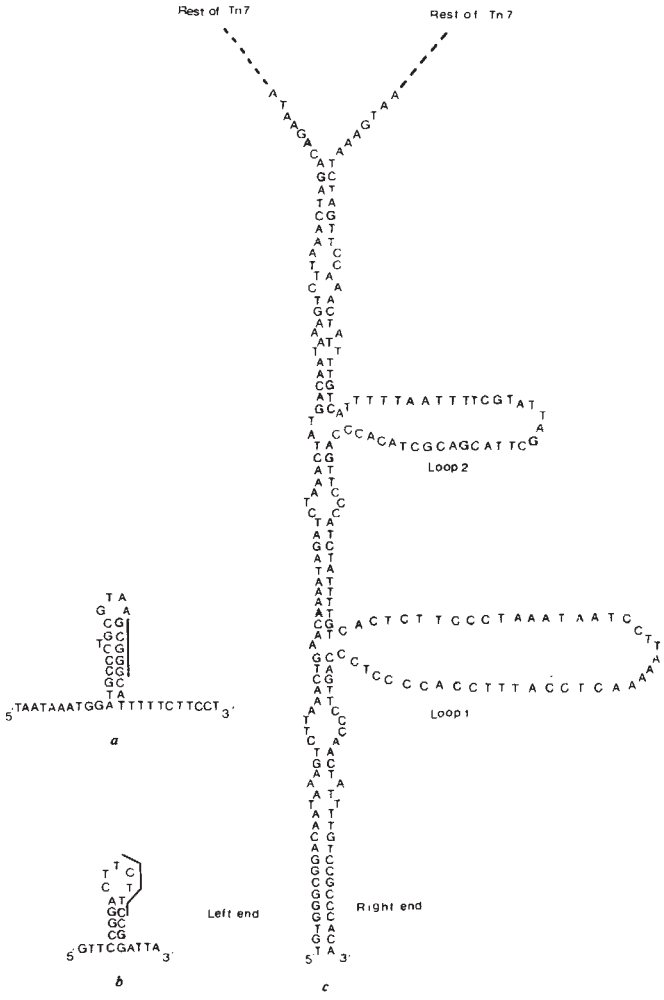


Fig. 3 *a*, The DNA sequence around the Tn*7* attachment site in the plasmid pCPL6 drawn as a potential hairpin structure that might be formed by intra-strand base pairing. This region is flanked by A+T-rich regions. Only one strand is indicated but the other would be present in a mirror image to form a cruciform. The bar marks the pentanucleotide sequence that is duplicated after Tn*7* insertion. *b* Shows a similar, but presumably less stable structure that might form at the Tn*7* insertion site in the plasmid ColE1. *c* Indicates a possible hairpin structure that might form by intra-strand base pairing at the ends of Tn*7*, when boxes 2 and 3 (as indicated in fig. 2*a*, *b*) are brought into register. Loops 1 and 2 are the intervening DNA sequences between the three box 3s in the right end of Tn*7*.

We thank R. P. Anderson and A. Otsuka for comments on the manuscript. The Tn*7* insertion was a gift from M. Saul.

Received 10 July 1981; accepted 28 April 1982.

1. Barth, P. T., Datta, N., Hedges, R. W. & Grinter, N. J. *J. Bact.* **125**, 800–810 (1976).
2. Lichtenstein, C. & Brenner, S. *Molec. gen. Genet.* **183**, 380–387 (1981).
3. Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
4. Bachmann, B. S., Brooks-Low, K. & Taylor, A. L. *Bact. Rev.* **40**, 116–167 (1976).
5. Fennewald, M. A. & Shapiro, J. A. *J. Bact.* **139**, 264–269 (1979).
6. Phillips, G. & Novick, R. P. *Nature* **278**, 476–478 (1979).
7. Bolivar, F. *et al. Gene* **2**, 95–113 (1977).
8. Miki, T., Hiraga, S., Nagata, T. & Yura, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5099–5103 (1978).
9. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
10. Tinoco, I., Uhlenbeck, O. C. & Levine, M. D. *Nature* **230**, 362–367 (1971).
11. Lilley, D. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6468–6472 (1980).
12. Panayototas, N. & Wells, R. D. *Nature* **289**, 466–470 (1981).
13. Gellert, M., Mizuuchi, K., O'Dea, M. & Nash, A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3872–3876 (1976).
14. Worcel, A. & Burgi, E. *J. molec. Biol.* **71**, 127–147 (1972).
15. Temin, H. *Cell* **21**, 599–600 (1980).
16. Zieg, J. & Simon, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4196–4200 (1980).
17. Hubermann, P., Klaer, R., Kuhn, S. & Starlinger, P. *Molec. gen. Genet.* **175**, 1127–1134 (1979).
18. Rothstein, R. J., Lan, L. F., Bahl, C. P., Narang, S. A. & Wu, R. *Meth. Enzym.* **68**, 98–109 (1979).
19. Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
20. Bentley, D. L. & Rabbitts, T. H. *Nature* **288**, 730–733 (1980).
21. Schreier, P. H. & Cortese, R. J. *J. molec. Biol.* **129**, 169–172 (1979).
22. Tomizawa, J-I., Ohmori, H. & Bird, R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1865–1869 (1977).
23. Staden, R. *Nucleic Acids Res.* **5**, 1013–1015 (1978).

Structural analysis of the prolactin gene suggests a separate origin for its 5′ end

Nancy E. Cooke* & John D. Baxter

Howard Hughes Medical Institute Laboratory, Department of Medicine, Biochemistry and Biophysics; and Metabolic Research Unit, University of California, San Francisco, California 94143, USA

Prolactin, growth hormone (GH) and placental lactogen constitute a family of polypeptide hormones which share certain structural[1–7] and functional[8] features. The prolactin and GH genes evolved from a gene duplication about 400 million years ago[9] and, in humans, segregated onto chromosomes 6 and 17 respectively[10,11]. We expand here previous reports[12,13] on the general organization of the rat prolactin gene, *rPRL*, assign its haploid gene number and extend its DNA sequence analysis. We postulate that a direct repeat flanking exon I may be the remnant of an insertional event.

A single plaque hybridizing to a *rPRL* cDNA[9] probe was isolated after screening 500,000 plaques of a λCharon4A library of Hooded rat chromosomal DNA. Digestion of the isolated recombinant phage DNA with *Pst*I or *Hin*dIII endonuclease resulted, in each case, in two fragments that hybridized to the *rPRL* cDNA probe (Fig. 1, lanes 3, 5). The four fragments were individually subcloned into pBR322 and used in conjunction with the recombinant phage to map and sequence *rPRL*. The single recombinant phage and its subclones encoded the

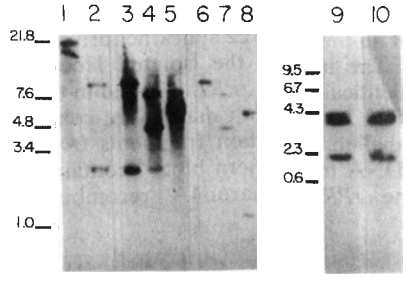


Fig. 1 Detetction of *rPRL* and distribution of coding and repetitive sequences among its restriction fragments. Lanes 1, 2: Sprague–Dawley rat liver DNA digested with *Bam*HI and *Pst*I respectively, electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose and hybridized to a ^{32}P-labelled *rPRL* cDNA probe[9]. Lanes 3–5: DNA fragments from the Hooded *rPRL* recombinant λCharon4A clone digested with *Pst*I, *Pvu*II and *Hin*dIII respectively, separated on a 0.8% agarose gel and hybridized to the ^{32}P-labelled *rPRL* cDNA probe. Lanes 6–8: as for lanes 3–5, but hybridized to ^{32}P-labelled total rat DNA after melting off the previous probe. Repetitive and coding sequences are both present in the larger of the two *Pst*I fragments and in the 5.6 kbp *Hin*dIII fragment, but the 1.5 kbp *Hin*dIII fragment contains only repetitive DNA and is located in the 3′-flanking region of the gene. Lanes 9, 10: autoradiograms of an identical blot of *Eco*RI-digested *rPRL* subclone pRPRL–*Pst*Q (see Fig. 2*c*) hybridized to ^{32}P-labelled rat DNA and then to a probe made from the 3′-repetitive region respectively. The positions of molecular size markers are indicated. Templates for probes were labelled internally with ^{32}P-dNTPs to a specific activity of 1×10^8 c.p.m. μg^{-1} using random primers[16] or nick translation[29]. Total DNA probes used to map repetitive sequences were synthesized by internally labelling rat DNA sheared to 500 bp. These probes have been shown to detect sequences in >50–100 copies per haploid genome in the hybridization conditions used[30]. *rPRL* was isolated from a Hoodel rat λCharon4A library (kindly provided by G. Paige) composed of partial *Eco*RI-digested rat DNA fragments ligated into phage arms. Established techniques were used for clone identification[31] and amplification[32]. Restriction endonucleases were purchased from New England Biolabs, Boeringer-Mannheim and Bethesda Research Laboratories, and used as recommended.

* Present address: Endocrine Section, Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

0028–0836/82/240603—04$01.00

© 1982 Macmillan Journals Ltd

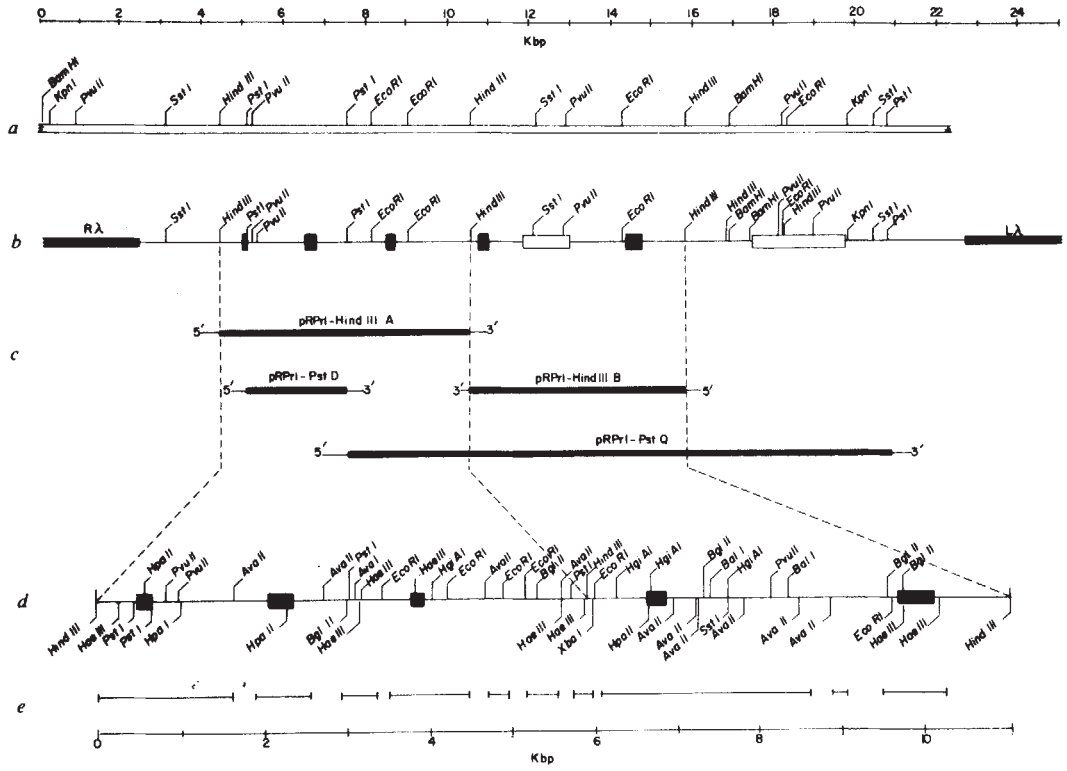


Fig. 2 Organization of *rPRL*. Restriction maps of *a*, the prolactin gene in Sprague–Dawley DNA; and *b*, the λCharon4A-*rPRL* recombinant clone. R and L refer to the right and left phage arms respectively, exons are indicated by solid rectangles and repetitive regions that hybridized to total DNA probes by white rectangles. *c*, Location and orientation in pBR322 of the λ-*rPRL* subclones used for detailed mapping and sequencing. *d*, Detailed structural map of the area sequenced and *e*, the extent of sequencing, indicated by the bracketed lines. *a–c* are drawn to the kbp scale on the top, and *d* and *e* to the scale on the bottom. The subclones indicated in *c* were constructed by ligating *Pst*I or *Hin*dIII digests of the λCharon4A–*PRL* recombinant into 5′-dephosphorylated (bacterial alkaline phosphatase, P-L Biochemicals) *Pst*I-or *Hin*dIII-digested pBR322 using T4 DNA ligase (New England Biolabs). *Escherichia coli* HB 101 were made competent by $CaCl_2$ treatment[33], transformed with the ligation mixtures, and recombinant colonies containing *rPRL* restriction fragments identified by hybridization to ^{32}P-labelled *rPRL* cDNA[34].

entire *rPRL* gene (Fig. 2*b*). The location, size and orientation of all the pBR322 subclones are shown in Fig. 2*c*.

The map of the isolated Hooded rat prolactin gene was identical to the *rPRL* mapped in the Sprague–Dawley genome (Fig. 2*a*, *b*). Both maps were constructed by hybridizing ^{32}P-labelled restriction fragments from specific regions of *rPRL* cDNA to Southern blots[14] containing restriction fragments of the *rPRL*–λCharon4A recombinant clone, or digests of Sprague–Dawley liver DNA. The 10.5 kilobase pair (kbp) *rPRL* was found to consist of five exons and four introns (Fig. 2*b*, *d*), similar to the pattern found in the rat GH gene (*rGH*)[15,16], and in good agreement with the size estimates predicted by heteroduplex mapping of the Sprague–Dawley *rPRL*[17]. The sizes of the *Pst*I genomic *rPRL* fragments and those from the *rPRL* recombinant phage DNA were identical (Fig. 1, lanes 2, 3). There were no *Bam*HI sites in the cloned gene, and only one hybridizing *Bam*HI genomic fragment was seen (Fig. 1, lane 1). Previously only one *rGH* was mapped in the rat genome[15,16], although multiple GH genes are present in the human genome[18]. It seems reasonable to conclude that there is only one *rPRL*, although the presence of additional prolactin genes embedded in an identical surrounding DNA environment is possible. A previously mapped and partially sequenced Sprague–Dawley *rPRL* lacks 2.55 kbp of DNA in intron D[12,13]. We found no evidence for this structure and conclude that it may represent a polymorphic variant, cloning artifact, or mapping error.

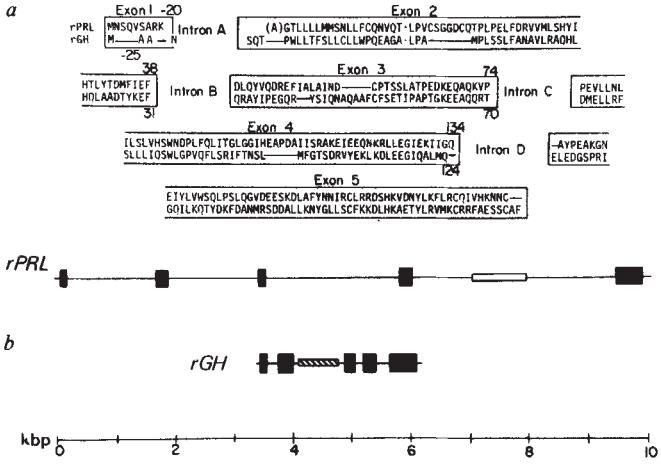


Fig. 3 Comparison of splicing site location and gene structure of *rPRL* and *rGH*. *a*, Amino acid sequences (single-letter code) of rat preprolactin and rat pregrowth hormone aligned to maximize homology[9]. Rectangles enclose the amino acids encoded by each exon. The amino acids located at the carboxy terminus of the coding region of each exon are numbered. In *rPRL*, the codons for alanine (A) or glycine (G) indicated at the beginning of exon 2 are actually split by intron A (see text for discussion). *b*, Comparison of the sizes of *rPRL* and *rGH*. Exons are indicated by solid rectangles, dispersed repetitive sequences by a white rectangle in *RPRL* and a hatched rectangle in *rGH*. Both genes are shown to the scale indicated at the bottom of the figure.

DNA sequence analysis has shown that the slicing junctions f introns A, B, C and D of *rPRL* are of the same class[19] and occur at analogous positions to those determined for *rGH* when both coding structures are aligned to maximize homology (Fig. 3*a*). Conservation of intron location in coding regions is frequently seen among individual members of gene families[19,20]. The conservation of splicing sites was particularly striking between *rPRL* and *rGH* as they retain only about 25% amino acid homology and *rPRL* is five times larger than *rGH* (Fig. 3*b*). This conservation of splicing site locations substantiates the hypothesis that *rPRL* and *rGH* evolved from a common precursor.

Although the similar overall structures of *rPRL* and *rGH* implied a common precursor, lack of homology and dissimilar lengths between exons I of these genes suggested that the 5′ ends may have had independent origins. A direct decanucleotide repeat which flanks exon I of *rPRL*, occurs at bases −77 to −68 in the 5′-flanking region and again in intron A at bases 340–349 (boxed in Figs 4, 5*a*). Similarly, an 80% homologous

© 1982 Nature Publishing Group

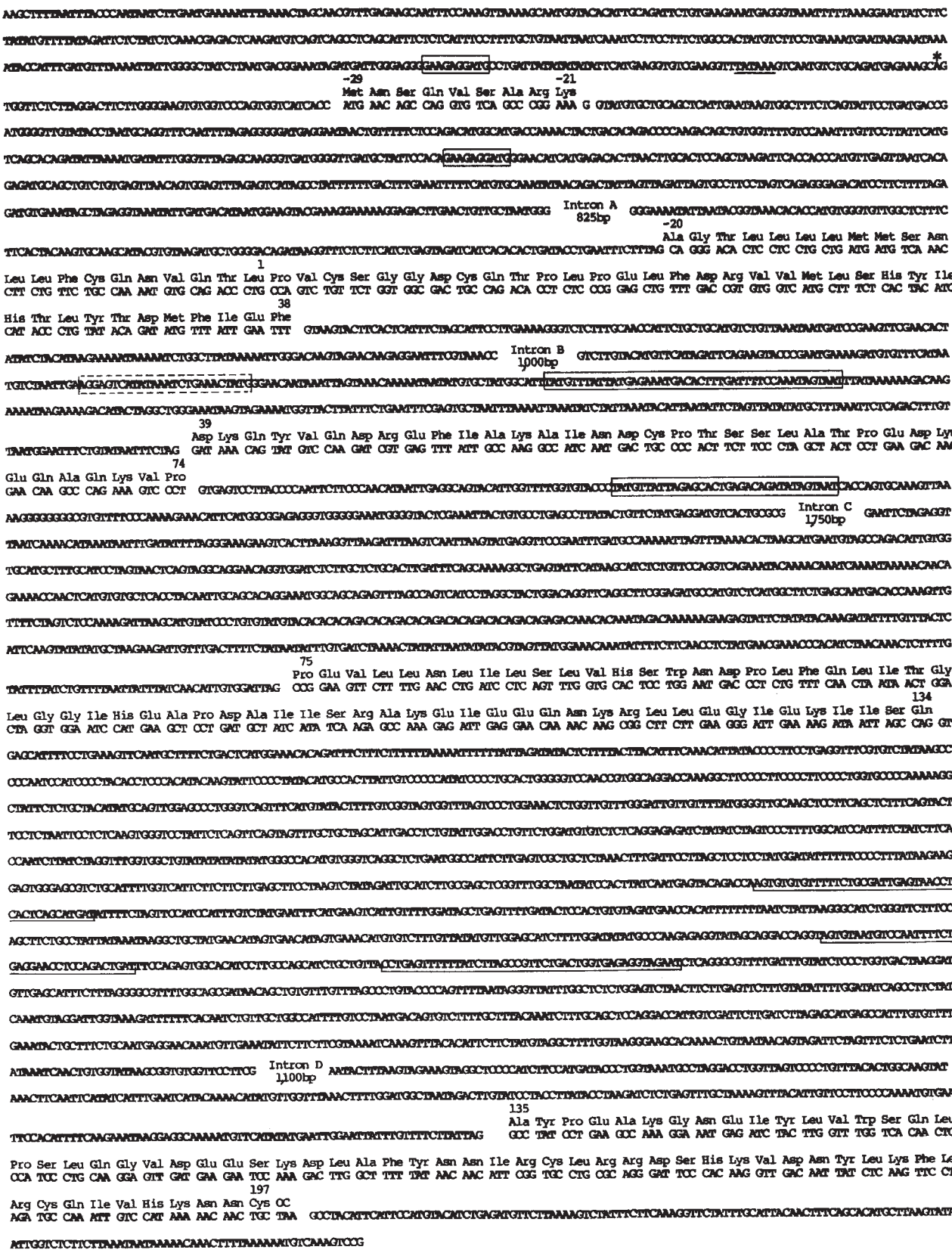


Fig. 4 DNA sequence of *rPRL*. The sequence begins 423 bp 5′ to the cap site, includes 5,070 bp of the gene and ends near the poly(A)addition site. The numbered amino acids of rat preprolactin are shown. The TATAAA sequence is underlined once, the putative capping site is starred. Two sets of direct repeats are enclosed in rectangles. A dashed rectangle encloses an area of homology with intron B of *rGH*. The dots in intron D demarcate the approximate limits of hybridization to total rat DNA probes and three areas of homology to human Alu family consensus sequences are underlined. The lengths of areas not presented are indicated. DNA sequencing was carried out by the chemical cleavage technique of Maxam and Gilbert[35], as previously modified[9]. Approximately 90% of the sequences shown were either determined twice or on both strands. DNA sequences were analysed in a PDP/1160 computer (Digital Equipment Corp.) using programs provided by H. Martinez[36].

© 1982 Nature Publishing Group

Fig. 5 Direct repeats flanking exon I of *rPRL*, *rGH* and conalbumin gene. *a*, Diagram of an isolated segment of the 5′-flanking region, promoter site, first exon and first intron of three genes. Rectangles enclose a decanucleotide repeat (100% homology) in *rPRL*, a 12 bp repeat (one mismatch, one insertion/deletion) in *rGH*[16] and a 14 bp repeat (79% homology, two insertions/deletions) in conalbumin gene[23]. The number of base pairs separating the 5′ repeat and TATAAA, TATAAA and the 5′ border of exon I, and the 3′ border of exon I and the 3′ repeat are indicated. The actual base number of the repeat units (cap site is +1) are indicated above or below the appropriate nucleotides. *b*, Shows an area of homology in the 5′-flanking regions of *rPRL* and *rGH*. The 5′ direct repeat in each gene, shown in *a*, is enclosed by a rectangle in *b*.

12 basepair (bp) repeat, different in sequence from the *rPRL* repeat, flanks exon I of *rGH* (Fig. 5*a*), occurring at bases −61 to −50 and 124 to 136[14]. In both genes, the repeats encompass TATAAA (a probable promoter region), the mRNA cap sites and exon I. The presence of these direct repeats may indicate that the encompassed sequence was inserted into these genes by mechanisms proposed for prokaryotic, viral and eukaryotic insertion sequences and transposons[21]. The insertion of mobile genetic elements characteristically results in a duplication of host target site DNA bracketing the inserted DNA. Although the sequences of *rPRL* and *rGH* are divergent, the DNA sequences encompassing the more 5′ of the direct repeats flanking exon I of r*PRL* and *rGH* were highly homologous (76%) (Fig. 5*b*). In both genes this homologous region was located in the approximate area that, in many other eukaryotic genes[22], contains the CAAT homology unit, which is lacking in both *rPRL* and *rGH*. Sixty-nine bp upstream from the 5′ copy of the decanucleotide direct repeat in *rPRL* was a second and presumably non-functional promoter sequence (AATAAA) preceded 42 bp upstream by a CAAAT sequence. *rGH* also has a second AATAAA, 21 bp 5′ to its repeat, but lacks the upstream CAAT sequence. These data suggest that the proposed insertion of DNA encoding exon I and its 5′-flanking region occurred just 3′ to the original promoter sequences of these genes in a homologous 'hotspot'[21] for integration. The proposed insertion of exon I may not be unique to *rGH* and *rPRL*. For example, a direct repeat flanking exon I in the chick conalbumin gene[23] was detected (Fig. 5*a*). As in *rPRL* and *rGH*, the direct repeats encompass the TATAAA sequence and the entire coding unit of the first exon. As exons I of *rPRL* and *rGH* are non-homologous and of different lengths, we concluded that they may have been inserted independently into already distinct prolactin and GH genes. Such an event could place new regulatory signals 5′ to these genes and facilitate divergence of regulation and function. The 73% nucleotide homology between the signal peptides of human and rat prolactin[7] suggested that the insertion of exon I, which partially encodes this region of the protein, occurred before the divergence of rat and human. Conservation of the decanucleotide repeats during further evolution may have been a random event or may indicate some functional constraint on their structure.

Introns A, B, C and D interrupt the coding regions of *rPRL* between codons −20/−19, 38/39, 74/75 and 134/135 respectively (Figs 3, 4). The precise positions of the splicing sites were assigned by locating the characteristic G–T at the 5′ end of the intron and A–G at the 3′ end of the intron[24], as well as by optimizing homology to U1 RNA[25–27]. The actual position of the 3′-splicing junction of intron A is ambiguous[13]. There are two possible splicing sites, one that includes an alanine (GCA) in the signal peptide (Fig. 4), and another that excludes this alanine and splices directly to glycine (GGG) at amino acid position −19. The former preserves the greatest complementarity with rat U1 RNA (60% compared with 40%). Two separate *rPRL* cDNAs sequenced from oestrogen-stimulated rats contained the alanine codon[12,28], while a third *rPRL* cDNA clone isolated from rat pituitaries maximally stimulated by oestrogen injections plus hypothalmic ablation did not contain this codon[9]. The identification of both mRNA species is consistent with the prediction of alternative splicing of the single *rPRL* gene. No similar situation occurs in the *rGH*. Whether the anomalous nature of this splice site relates to a separate evolutionary history of exon I in *rPRL* and *rGH* remains speculative.

Two regions of dispersed repetitive sequences were present in the cloned 20 kbp genomic fragment containing *rPRL*. Their presence and location were determined by Southern blot hybridization using ^{32}P-labelled total rat DNA as probe (Fig. 1, lanes 6–9). Results of more detailed mapping localized these two repetitive DNA sequences to intron D and to the 3′-flanking region (Fig. 2*b*). These two repetitive regions do not hybridize to each other, or to the repetitive DNA in intron B of *rGH*. There are therefore at least three discrete families of dispersed repetitive elements in the rat genome represented in and around these two related genes. Insertion of these repetitive sequences may serve rapidly to expand intron DNA and may limit interaction between the two genes, thus facilitating their structural divergence.

We thank Ilana Mittman for technical assistance, and Drs Stephen Liebhaber, John Shine and Y. W. Kan for advice, assistance and support.

Received 19 October 1981; accepted 14 April 1982.

1. Catt, K. J., Moffat, B. & Niall, H. D. *Science* **157**, 321 (1976).
2. Li, C. H., Dixon, J. S., Lo, T. B., Pankov, Y. M. & Schmidt, K. D. *Nature* **224**, 695 (1969).
3. Sherwood, L. M. *Proc. natn. Acad. Sci. U.S.A.* **58**, 2307–2314 (1967).
4. Niall, H. D., Hogan, M. L., Sauer, R., Rosenblum, I. Y. & Greenwood, F. C. *Proc. natn. Acad. Sci. U.S.A.* **68**, 866–869 (1971).
5. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 486–494 (1977).
6. Shine, J., Seeburg, P. H., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 494–499 (1977).
7. Cooke, N. E., Coit, D., Shine, J., Baxter, J. D. & Martial, J. A. *J. biol. Chem.* **256**, 4007–4016 (1981).
8. Niall, H. D. *et al. Recent Prog. Horm. Res.* **29**, 387–416 (1973).
9. Cooke, N. E., Coit, D., Weiner, R. I., Baxter, J. D. & Martial, J. A. *J. biol. Chem.* **255**, 6502–6510 (1980).
10. Owerbach, D., Rutter, W. J., Cooke, N. E., Martial, J. A. & Shows, T. B. *Science* **212**, 815–816 (1981).
11. Owerbach, D., Rutter, W. J., Martial, J. A., Baxter, J. D. & Shows, T. B. *Science* **209**, 289–291 (1980).
12. Gubbin, E. J., Maurer, R. A., Legrimini, M., Erwin, C. R. & Donelson, J. E. *J. biol. Chem.* **255**, 8655–8662 (1980).
13. Maurer, R. A., Erwin, C. R. & Donelson, J. E. *J. biol. Chem.* **256**, 10524–10528 (1981).
14. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
15. Page, G. S., Smith, S. & Goodman, H. M. *Nucleic Acids Res.* **9**, 2087–2104 (1981).
16. Barta, A., Richards, R., Baxter, J. D. & Shine, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4867–4871 (1981).
17. Chien, Y. & Thompson, E. B. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4583–4587 (1980).
18. Goodman, H. M. *et al.* in *Mobilization and Reassembly of Genetic Information* (eds Scott, W. A., Werner, R., Joseph, D. R. & Schultz, J.) 155–179 (Academic, New York, 1980).
19. Sharp, P. A. *Cell* **23**, 643–646 (1981).
20. Efstratiadis, A. *et al. Cell* **21**, 653–668 (1980).
21. Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
22. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. *Nucleic Acids Res.* **8**, 127–142 (1980).
23. Cochet, M. *et al. Nature* **282**, 567–574 (1979).
24. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
25. Seif, I., Khoury, G. & Shar, R. *Nucleic Acids Res.* **6**, 3387–3398 (1979).
26. Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877–1879 (1980).
27. Lerner, M. Boyle, J., Mount, S., Wolin, S. & Steitz, J. *Nature* **283**, 220–224 (1980).
28. Taylor, W. L., Collier, K. J., Weith, H. L. & Dixon, J. E. *Biochem. biophys. Res. Commun.* **102**, 1071–1077 (1981).
29. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
30. Shen, C. J. & Maniatis, T. *Cell* **19**, 379–391 (1980).
31. Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
32. Maniatis, T. *et al. Cell* **15**, 687–701 (1978).
33. Dagert, M. & Ehrlick, D. S. *Gene* **6**, 23–28 (1979).
35. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
35. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
36. Martinez, H. M., Bock, S. A. & Katzung, B. *DNA Programs Manual* (Department of Biochemistry and Biophysics, University of California, San Francisco, 1980).

© 1982 Nature Publishing Group

BOOK REVIEWS

The bones of palaeoanthropology

Andrew Hill

SOMETIMES in science a totally new class of things becomes a legitimate object of study, not through having previously been undiscovered, but through previously mundane items acquiring relevance. Recently, palaeontologists and archaeologists formed an interest in natural accumulations of modern broken bones as an aid to understanding the origins and significance of vertebrate fossil assemblages. The question of what happens to animals between death and possible fossilization has even been blessed with the status of its own name — taphonomy — although it is little more than a necessary prior consideration of any reliable palaeoecological work. A number of fairly discrete centres of interest have coalesced to constitute the activity as it is developing today, and one of the most important of these relates to Plio-Pleistocene hominids in sub-Saharan Africa. Amongst scientists involved, C.K. Brain stands out as the pioneer; this impressive book is a statement of his investigations.

Before the discovery of many early hominids in eastern Africa most of the evidence for human evolution came from a few caves in the far south. The hominids and associated bone accumulations became controversial, mainly because Raymond Dart plausibly suggested most of the bones were tools and weapons fashioned by australopithecines. He also claimed evidence of violence, offered not only to other species but to other hominids. This is the aggressive view propagated by Robert Ardrey. But such ignoble savagery is no longer in vogue and nature's teeth and claws are less red than they were. Beware of the dogma; could it instead have been cave carnivores that produced the bone remains in feeding, or was it really the hominids that were responsible for — as Philip Tobias put it — the "lion's share"? This is an important matter relevant to our sense of the behaviour of early man, and maybe to our feelings about our own nature today. It is around this question that *The Hunters or the Hunted?* is structured.

The Hunters or the Hunted? An Introduction to African Cave Taphonomy. By C.K. Brain. Pp.365. ISBN 0-226-07089-1. (University of Chicago Press: 1981.) £24.50, $45.50.

Brain attempts to provide an answer by looking in some detail at a large spectrum of contemporary creatures that accumulate bones, both in caves and in more open situations. Consequently the first part of the book is devoted to accounts of the extensive observations and the results of experiments on the feeding and food remains of modern human beings and their dogs, and of cave carnivores such as leopards and hyaenas. Porcupines are curious collectors of bones which are also dealt with, as are birds such as eagles and owls which contribute quantities of small vertebrate material to cave sediments. In addition there is a discussion of the middens of later prehistoric human cave dwellers in southern Africa. A number of aspects of the accumulations are investigated, such as the various kinds of bone breakage produced by these different factors, the size and numbers of the prey, and the proportional representation of different skeletal parts in the collections. Based upon such features as these Brain successfully establishes criteria for distinguishing between most of the collecting agencies.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

A stylized impression of the Swartkrans leopard hypothesis, a theory which might account for the abundance of hominid remains in the Swartkrans cave. An australopithecine is being consumed in a *Celtis* tree overhanging a cave, while hyaenas of several species wait hopefully below. (Picture courtesy of C.K. Brain.)

Next follows an incidental but most useful summary of information about all fossil animal species from the Sterkfontein Valley caves. The last section takes each of these cave systems in turn — Sterkfontein, Swartkrans and Kromdraai — and within them each stratigraphic member, in order to assess the origins of the bone accumulations. In this the comparative information presented in Part One is employed. An extremely valuable element is a complete taxonomic listing of all the fossil vertebrate specimens from the deposits, and there is an historical description of the discovery and subsequent work at the sites. Notes are also provided on the other South African australopithecine caves — Taungs and Makapansgat — that Brain has not studied in such detail. All of this is thoroughly documented, supported by an appendix of 121 tables, and there are many illustrations including some unfortunately rather murky photographs.

New branches of activity in natural science have a tendency to reach their anecdotage before maturity. Is this a necessary stage of development, or an avoidable fault? It is certainly a feature of the first part of the book, where although many of the accounts add greatly to the interest, they also seem to overload the argument with trivial facts that add nothing substantial or lead nowhere. The data — like the

© 1982 Nature Publishing Group

food remains that Brain describes — are not fully digested, but they will provide a rich resource for the depredations of other scientific scavengers. It would be useful for some of these to be of a more statistical inclination. There are, admittedly, methodological difficulties with the analysis of the kind of numerical information that Brain has collected; conventional statistical manipulations are always possible, but it is not always clear that the answers mean much in palaeoecological terms. Certainly his tentative approach, drawing essentially qualitative and specific conclusions through closely argued reasoning, yet presenting the original numerical data in full detail, is better than making more suspicious assertions based on operations only superficially possessing greater mathematical rigour. However, a statistical attack on the problem might be beneficial, for I feel that the data contain more information than the author has at present extracted from them; not that his conclusions are slight.

Brain's meticulous reappraisal seems conclusively to show little evidence of early hominids flailing at each other and at other animals with bone clubs or other weapons. Most of the assemblages are postulated to be the product of hyaenas and large felids such as leopards and the extinct *Dinofelis* that controlled the caves, probably a result of opportunistic predation at hominid sleeping sites. But between Sterkfontein Members 4 and 5 came a change, corresponding to that crucial step in human evolution when man put the cat out for the first time. The later bone collection has an exclusively human character, showing that we had at last attained some dominance over other predators. Brain's work, however, supports the idea that initially the hominids relied heavily upon scavenging before increasing intelligence and the beginning of technology permitted the greater development of hunting abilities.

It is gratifying to see these very early hominids regarded as animals on their own terms, as species which are extinct and, no doubt, ones which had their own appropriate mode of life and behaviours. Identifying them too simply with models founded either on modern technologically undeveloped peoples or on modern pongids sometimes masks this fact, and it can be avoided by dealing with the matter from a more fundamental ecological basis, as Brain does here.

The Hunters or the Hunted? is a very important book for palaeoanthropology. It presents the first thorough analysis of the Sterkfontein Valley assemblages, contributes significantly to the resolution of lingering controversies and, by placing the old information in a fresh perspective, enables new and more sophisticated questions to be asked not only of the South African material but of similar assemblages elsewhere. Another contribution is that it reinforces the recent change in feelings as to what constitutes data, for the value of looking at fossil and contemporary bones as closely as this is clear. Brain urges the necessity of recovering fossils with a high regard for subtle detail. I hope that excavators of any vertebrate fossil site will be persuaded to follow his advice and pay more attention to these features of bone accumulations that have been previously neglected; for taphonomy can be a powerful tool in elucidating the problems of fossil assemblages, especially when handled with the care and caution that Brain brings to the subject. □

Andrew Hill is a Research Fellow at Harvard University.

Journals' review issue 1982

On October 7th *Nature* will publish a second review supplement devoted to science journals. Last year's supplement, covering journals first published between January 1978 and May 1980, appeared in *Nature* **293**, 341-369; see p.343 of that issue for details.

Criteria for inclusion of a journal in the 1982 issue are that:

- the first number appeared, or the journal was re-titled, between June 1980 and May 1981;
- it is published at least three times a year;
- the main language used is English.

Broadly, periodicals of professional interest to scientists will be considered for review, with the exception of abstracts' journals.

In addition it is hoped to cover publicly available newsletters, first published between January 1978 and May 1981, in that issue.

Publishers and societies are invited to submit four sample issues of periodicals satisfying the above criteria, *including the first and most recent numbers*, to the Review Editor, *Nature*, 4 Little Essex St, London WC2R 3LF, England (London 836 6633 ext 2570) as soon as possible.

Structure in (and of) the neurosciences

Margaret Boden

Meaning and Purpose in the Intact Brain: A Philosophical, Psychological, and Biological Account of Conscious Processes. By Robert Miller. Pp.239. ISBN 0-19-857579-3. (Clarendon/Oxford University Press: 1982.) £20, $34.50.

ROBERT Miller is a neuroscientist dissatisfied with neuroscience. It has not produced a general theory of consciousness, nor does it regard questions about consciousness, purpose, meaning, understanding or cognition as central. It has contributed little to our understanding of the behaviour of the whole organism, the overall patterns and strategies of living. Instead of "it" one should perhaps put "they", for there are a bewilderingly diverse and narrowly specialized set of neurosciences. In general, neuroscientists ignore psychology and deliberately eschew philosophical speculation.

One cannot but agree with Miller's criticisms of neuroscience. One may even applaud his aim of synthesizing brain research, psychology and philosophy in the search for a theory of consciousness. Nevertheless, despite its laudable intentions, this book is a failure. It does not help us progress significantly towards a theoretical understanding of meaning, purpose, cognition or consciousness.

This failure is grounded in Miller's reliance on an outmoded method of argument, and — ironically, given his critique of professional specialism — in his ignoring recent concepts and forms of argument which offer much greater promise of bringing neuroscience closer to psychology (and philosophy).

Characteristically, Miller argues from structure to function or vice versa. So he gives an account of consciousness in terms of "omniconnected networks", sets of neurones showing synaptic plasticity and influenced by transmitter substances of various kinds. Doubtless such networks are involved. But, as computer science has helped to show, the notion that specific cognitive functions can be attributed to specific cerebral mechanisms on structural grounds alone is mistaken.

Moreover, computational models of psychological processes provide concepts better suited to express specific cognitive functions than the vague psychological terminology used by Miller. Marr's physiologically-informed work on the visual system, for instance, is nowhere mentioned (though his earlier work on the neocortex is cited). Marr argued that neuroscientists will need to specify the computational tasks that the brain must (or might) be performing, before discovering which physiological mechanisms are performing them. Whether or not one regards this claim as overly sweeping, it is effecting radical changes in our view of physiological psychology. Miller's book is gravely flawed by its failure to consider the usefulness of computational ideas in neuroscience. □

Margaret Boden is Professor of Philosophy and Psychology at the University of Sussex. Her books include Purposive Explanation in Psychology *(Harvard University Press, 1972) and* Artificial Intelligence and Natural Man *(Basic Books, 1977).*

© 1982 Nature Publishing Group

The ore deposit—tectonics connection

Frederick J. Sawkins

Mineral Deposits and Global Tectonic Settings. By A.H.G. Mitchell and M.S. Garson. Pp.404. ISBN 0-12-499050-9. (Academic: 1982.) £23.60, $48.50.

WITH *Mineral Deposits and Global Tectonic Settings,* Mitchell and Garson are to be congratulated on providing us with a most impressive piece of scholarship. The interrelationships between tectonics and mineral resources have attracted increasing attention from earth scientists during the past decade, and their book represents by far the most competent and complete analysis of the subject which has emerged so far.

With commendable thoroughness the authors deal successively with hotspot and rifting environments, passive margins and continental margins, oceanic settings, subduction environments and collision settings, and finally transform faults, both in terms of their geology and their associated mineral deposits. The subject is approached in a no-frills, workmanlike style and few stones are left unturned, especially with respect to discussion of the various geological sub-environments within the major types of plate boundary settings. The 164 illustrations that accompany the text are, in the main, well chosen and illuminating, and the bibliography is not only extensive, but consists mainly of papers published within the past decade. The treatment is truly world-wide in scope, although the number of examples of geological terrains and mineral deposits taken from Communist-bloc countries is understandably rather small.

In view of the foregoing, it is perhaps less than charitable to focus on what I regard as the relatively minor shortcomings of the volume. The great challenge inherent in such an undertaking is the achievement of balance, presumably guided, at least in part, by economic significance. Here the authors' greater experience in Africa and the Far East, as opposed to the Cordillera of North and South America, is discernible. For example, considerable emphasis is placed on alkaline complexes and ophiolites, and important Cordilleran deposits such as polymetallic skarns and epithermal precious metal deposits are given short shrift. Porphyry copper deposits are also dealt with in surprisingly few pages considering their pre-eminent economic importance and well-defined tectonic setting. Also, although tonnage and grade figures are given for some major deposits, the distinction between mere mineral occurrences, small deposits and major orebodies is not at all clear in many instances.

This book, however, represents mandatory reading for anyone interested in mineral deposits–tectonics relationships. It will be a valuable resource to those concerned with any aspect of earth resources, tectonics or, for that matter, the regional geology of our planet. □

Frederick J. Sawkins is a Professor in the Department of Geology and Geophysics at the University of Minnesota.

Plants in the sea

G.E. Fogg

Marine Botany. By Clinton J. Dawes. Pp.628. ISBN 0-471-07844-1. (Wiley: 1981.) £33.50, $59.85. *The Biology of Seaweeds.* Edited by Christopher S. Lobban and Michael J. Wynne. Pp.786. UK ISBN 0-632-00672-2; US ISBN 0-520-04585-8. (Blackwell Scientific/University of California Press: 1982.) £45, $85.

THE plants of the sea are phylogenetically very different from their terrestrial counterparts, and in open waters competitive advantage is with the small rather than with the large and complex as on land. Such differences justify treating marine botany as a distinct discipline but the task of surveying the whole field, including the taxonomy, life histories, biochemistry, physiology and ecology of marine plants, has not often been attempted. However, with the growth of our knowledge of the sea and the extension of our ability to see and handle marine plants *in situ* provided by SCUBA diving techniques, it is becoming easier to shed preconceived ideas derived from land-based botany and discern general principles.

Marine Botany is such a survey, designed chiefly for the undergraduate student. The author has set out to cover physiology and biochemistry, as well as taxonomic and ecological aspects, but the treatment is mainly descriptive. Thus a conventional account of the various groups of marine algae, of a sort which is available in many other texts, takes up almost half the book. The space left for other topics is not always used to best advantage — a picture of an oxygen meter is not very informative. The single chapter on plankton says nothing of one of its most fascinating aspects — the interplay of buoyancy control and exploitation of water movements which enables this type of organism to live a precarious existence between the darkness of the depths and nutrient starvation at the surface.

Rather better are the chapters on seagrass communities, mangrove swamps and coral reefs, which give accounts of communities not often featured in botanical texts, and that on economic utilization is an up-to-date summary. There are appendices on marine fungi and bacteria and these, too, contain material which is not readily available elsewhere. A feature of the book is that details of chemical and other techniques and keys for identification for some groups as far as order or genus are given. The author has written a book which will be useful as an introduction, but has rather missed the opportunity to display the beautiful adaptations in form and function of plants to the marine environment.

The Biology of Seaweeds, a multi-author work, is more limited in scope and much more advanced in treatment. It is organized in four sections — on structure and reproduction, ecology, physiology and biochemistry, and seaweeds as resources. The first of these summarizes recent advances in our knowledge of red, brown and green algae. For those already bemused by the intricacies of the life-cycles of red seaweeds these chapters will not help; but that physiologists, biochemists and ecologists must take account of such things is shown in an excellent account of polysaccharides — in which alternation of generations at the biochemical level is described — by the well-supported suggestion that having alternate life-forms better adapts a seaweed to environmental stress, and by casual observations such as that the crustose phase of *Scytosiphon lomentaria* is less palatable to the snail *Littorina littorea* than is the foliose plant.

Better understanding of life-cycles has arisen largely from increasing use of laboratory cultures, and cultures are also being used extensively in studies on nutrient uptake, morphogenesis, growth regulators and other biochemical and physiological aspects, while transplantation and other *in situ* manipulations are now standard techniques for the field ecologist. Paralleling this is the development of methods in mariculture to meet the growing demand for seaweed products in industry. Chapters on geographical distribution, populations, various aspects of seaweed photosynthesis, translocation, chemical constituents and sexuality are all authoritative and thought-provoking. A statement that a water movement of 14 m sec^{-1} exerts a stress equivalent to that of a gale of 890 miles per hour on land makes one wish that more had been said about the mechanical properties of seaweeds, in particular their relation to the molecular architecture of the cell wall.

However, the study of seaweeds is expanding too rapidly for any survey to be comprehensive. Apart from the proceedings of the ten international seaweed symposia which have now been held, this book is the first major work to bring together knowledge of the various aspects of these marine plants under one cover. It is an important sea-mark in marine botany that will soon be left astern. □

G.E. Fogg is Professor of Marine Biology at the University College of North Wales, Bangor.

© 1982 Nature Publishing Group

A microbiologist's vade-mecum

John Postgate

The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria. Edited by Mortimer P. Starr *et al.* Two parts, pp.2,596, ISBN 3-540-08871-7. (Springer-Verlag: 1982.) DM 880, $410.

TO DO this magnificent work justice would require a team of reviewers almost as numerous as its 182 contributors. For it is nothing less than a comprehensive handbook of bacteriology: a detailed and up-to-date guide to the isolation, cultivation and identification of almost all the known bacteria.

The two volumes are divided into 23 sections (A–W inclusive), each of which comprises between 2 and 18 chapters. Section A consists of 6 introductory essays, the first of which is an excellent account of prokaryotic diversity which contrives to be highly informative as well as entertaining — see Table 13, for example. The other essays cover habitats, including diseased man and animals, as well as plants; principles of isolation, cultivation and conservation; and principles of characterization and identification. I found each of them thorough and informative; they provide an excellent *hors d'oeuvre* to the subsequent 22 courses.

I shall not have space to review those in any depth, so I shall indicate the pattern of the books by describing Section B, which deals with the phototrophic bacteria. It consists of 12 chapters. After a general introduction presenting their physiology and anatomy, come highly practical chapters on the isolation of cyanobacteria from normal and saline environments. A contribution on the habitats, isolation and characterization of vacuolate cyanobacteria is followed by one of broadly similar pattern on thermophilic cyanobacteria. Next is a more taxonomic contribution on the genera of cyanobacteria. The prochlorophytes are then briefly dealt with, preceding a short account of the isolation and cultivation of rhodospirilla. The genus *Ectothiorhodospira* gets its own chapter, followed by the chromatia and chlorobia; these two chapters collectively cover the phototrophic sulphur bacteria. The isolation of the photosynthetic flexibacteria is discussed next, and the section is concluded by a key to the classification of the non-oxygenic phototrophs. As this description indicates, the phototrophic bacteria are dealt with very thoroughly and the emphasis is generally highly practical: on media, culture, enrichment, storage and so on.

This pattern, a general essay introducing the detailed articles, is more or less followed in each section throughout the book. Section C is on the gliding bacteria, and D — the shortest, containing two chapters — deals with some sheathed and vacuolate bacteria. Section E covers budding and appendaged bacteria; F, the spirochaetes; G, some helical and curved bacteria; H, the pseudomonads; I, the diazotrophs with *Agrobacterium* and *Alcaligenes* thrown in. Section J is a ragbag of mixed chemotrophs (sulphate reducers, methylotrophs, methanogens, halobacteria and so on), the least logical but nevertheless a most useful section. Section K covers obligate chemotrophs, bringing us to the end of what might broadly be called "environmental" bacteria.

It is a pity that Section L, the last in Vol.1, covering medically important Gram-negative rods (*Brucella, Legionella* and so on), was not transferred to the second volume but I suppose binding problems would have arisen. Volume 2 itself covers *inter alia* most of the medically important bacteria, as well as rumen and spoilage organisms. Section M is on enterobacteria and vibrios; N, various facultative rods; O, anaerobic rods; P and Q, cocci; R and S, spore formers. Section T, on coryneforms, leads logically to Section U on actinomycetes. Finally, Section V deals with the obligate symbionts and W with the wall-deficient bacteria.

I found it impracticable to go through all 169 chapters in detail but I have read carefully about 20 that I know something about and have dipped into most of the rest. As with the introductory essays, I am impressed by the quality and conciseness of the writing and the high information content. The authors are leading experts in their particular fields and their names, too many to list, will be familiar to all microbiologists. Some of the arrangement is idiosyncratic, however, and some of the classifications are inconsistent with the eighth edition of *Bergey's Manual* (Williams & Wilkins, 1974), usually for good reasons. No handbook of this kind can be perfect: I combed Section C (gliding bacteria) for a cross-reference to *Desulfonema,* which rightly appears among sulphate-reducing bacteria, although it glides. It is also a pity that all the colour pictures are clustered at the front of Vol.1.

A major concern to most potential customers will be the overlap with *Bergey's Manual.* Certainly, *The Prokaryotes* covers much of the information available in the eighth edition of *Bergey* but it is a far more ambitious and up-to-date work: it is not simply a determinative system, but a real laboratory handbook for microbiologists. I recommend it strongly to any microbiological library; it is an invaluable guide and reference work which ought to be readily available to every working microbiologist. □

John Postgate is Professor of Microbiology at the University of Sussex and Director of the Agricultural Research Council's Unit of Nitrogen Fixation.

Ubiquitous gels

E.G. Richards

Electrophoresis: Theory, Techniques, and Biochemical and Clinical Applications. By Anthony T. Andrews. Pp.336. ISBN 0-19-854626-2. (Clarendon/Oxford University Press: 1982.) £25, $59.

DR Andrews remarks in his preface to this volume that well over half of all research papers in biochemistry involve electrophoresis. This impressive statistic calls for comment. Books with similar titles published 30 years ago would be mainly about free boundary electrophoresis à la Tiselius, a technique requiring complicated and expensive equipment; in contrast the present book is predominantly about zonal electrophoresis in polyacrylamide gels. For this the equipment is simple and cheap; the method is easy to master and no great sophistication is required to interpret the results. Anybody can do it, and, as the literature suggests, does.

Gels, polyacrylamide, starch or agarose provide the most favoured media for zonal electrophoresis with polyacrylamide well ahead of the field. It is 23 years since Raymond and Weintraub introduced the latter and thus strange that there is still no clear picture of the structure of such gels or how they exert their sieving effect in the electrophoretic process. Such understanding as we have is empirical. A consequence is that the technique still resembles more an art than a science. Recipes, with which the book abounds, are handed down from generation to generation of research students, and anomalies and artifacts, many of which Dr Andrews discusses, are there to dishearten the novice and mystify the expert.

Such theory that there is is adequately displayed, albeit occasionally a little shakily, and it is a minor criticism that it is developed in parallel for proteins and for nucleic acids. This leads to both repetition and a failure to indicate how lessons learnt from one class of macromolecules are relevant to the other.

Thus the major thrust is practical. Dr Andrews is to be congratulated on providing under one cover, succinct, clear and practical descriptions of electrophoresis in polyacrylamide and other media in most of the usual variations, as well as of related techniques including isoelectric-focusing, isotachophoresis, immuno-electrophoresis and affinity electrophoresis. □

E.G. Richards is Senior Lecturer at the Department of Biophysics at King's College, University of London.

© 1982 Nature Publishing Group

BOOKS RECEIVED

Biological Sciences

COMPER, W. D. Heparin (and Related Polysaccharides). Polymer Monographs, Vol. 7. Pp.254. ISBN 0-677-05040-2. (Gordon & Breach: 1981.) $55.95.

CONRAT-FRAENKEL, H. and KIMBALL, P. C. Virology. Pp.406. ISBN 0-13-942144-0. (Prentice-Hall: 1982.) $32.95.

COOPER, S. J. (ed.). Theory in Psychopharmacology, Vol. 1. Pp.338. ISBN 0-12-18800-1. (Academic: 1981.) £19.20, $39.50.

DALE, J. E. The Growth of Leaves. Studies in Biology, 137. Pp.57. ISBN 0-7131-2836-4. (Edward Arnold: 1982.) £2.25.

EISLER, R. Trace Metal Concentrations in Marine Organisms. Pp.685. ISBN 0-08-025975-8. (Pergamon: 1981.) $120.

FELDBERG, W. S. Fifty Years On: Looking Back on Some Developments in Neurohumoral Physiology. Pp.106. ISBN 0-85323-364-0. (Liverpool University Press: 1982.) £6.25.

GALLAGHER, R. H. *et al.* (eds). Finite Elements in Biomechanics. Pp.401. ISBN 0-471-09996-1. (J. Wiley & Sons: 1982.) £21, $49.50.

GUTHRIE, W. and FAWKES, R. A Colour Atlas of Surgical Pathology. Pp.220. ISBN 0-7234-0759-2. (Wolfe Medical: 1982.) £35.

HALL, J.C., GREENSPAN, R.J. and HARRIS, W.A. Genetic Neurobiology. Pp.284. ISBN 0-262-08111-3. (MIT Press: 1982.) $29.95.

HAWK, G.L. (ed.). Biological/Biomedical Applications of Liquid Chromatography III. Chromatographic Science, Vol.18. Pp.440. ISBN 0-8247-1297-8. (Marcel Dekker: 1981.) SwFr. 148.

HAY, J.G. and REID, J.G. The Anatomical and Mechanical Bases of Human Motion. Pp.443. ISBN 0-13-035139-3. (Prentice-Hall: 1982.) £14.95.

HAYAT, M.A. Fixation for Electron Microscopy. Pp.501. ISBN 0-12-333920-0. (Academic: 1981.) $50.

HECHT, M. K., WALLACE, B. and PRANCE, G. T. (eds). Evolutionary Biology, Vol. 14. Pp.437. ISBN 0-306-40775-2. (Plenum: 1982.) $39.50.

HÉLÈNE, C. *et al.* (eds). Trends in Photobiology. Proc. of the 8th International Congress on Photobiology and of the Colloque International du CNRS held July 1980, Strasbourg, France. Pp.673. ISBN 0-306-40644-6. (Plenum: 1982.) $35.

HENLE, W. *et al.* (eds). Current Topics in Microbiology and Immunology, Vol. 94/95. Pp.308. ISBN 3-540-10803-3. (Springer-Verlag: 1981.) DM 132, $61.50.

HENLE, W. *et al.* (eds). Current Topics in Microbiology and Immunology, Vol.97. Pp.204. ISBN 3-540-11118-2. (Springer-Verlag: 1982.) DM 94, $43.80.

HENRICH, H. (ed.). Microvascular Aspects of Spontaneous Hypertension. Pp.186. ISBN 3-456-81135-7. (Hans Huber: 1982.) SwFr.53, DM58.

HEYM, C. and FORSSMANN, W.-G. (eds). Techniques in Neuroanatomical Research. Pp.395. ISBN 3-540-10686-3. (Springer-Verlag: 1981.) DM 118, $55.

HOERNI-SIMON, G. and ZITTOUN, R. (coordinateurs) Progrés en Hématologie, 2. Décembre 1981. Les Lymphomes Malins Non Hodgkiniens. Pp.166. Pbk ISBN 2-7040-0406-4. (Doin Éditeurs, Paris: 1981.) Pbk np.

HOLZER, H. (ed.). Metabolic Interconversion of Enzymes, 1980. International Titisee Conference, October 1980. Pp.212. ISBN 3-540-10979-X. (Springer-Verlag: 1981.) DM 129, $60.10.

HOOK, E.B. and PORTER, I.H. (eds). Population and Biological Aspects of Human Mutation. Pp.435. ISBN 0-12-355440-3. (Academic: 1981.) $34.50.

HOPPENSTEADT, F. Mathematical Methods of Population Biology. Cambridge Studies in Mathematical Biology, 4. Pp.149. Hbk ISBN 0-521-23846-3; pbk ISBN 0-521-28256-X. (Cambridge University Press: 1982.) Hbk £15; £6.50.

HORECKER, B.L. and STADTMAN, E.R. (eds). Current Topics in Cellular Regulation, Vol.20 — 1981. Pp.317. ISBN 0-12-152820-0. (Academic: 1981.) $39.

HRABA, T. and HASEK, M. (eds). Cellular and Molecular Mechanisms of Immunologic Tolerance. Pp.600. ISBN 0-8247-1552-7. (Marcel Dekker: 1981.) SwFr. 178.

ILMAVIRTA, V., JONES, R.I. and PERSSON, P.-E. (eds). Lakes and Water Management. Developments in Hydrobiology, 7. Proceedings of the 30 Years Jubilee Symposium of the Finnish Limnological Society, held in Helsinki, Finland, September 1980. Pp.222. ISBN 90-6193-758-2. (W. Junk, The Hague: 1982.) Dfl.125, $54.50.

INGLE, D.J., GOODALE, M.A. and MANSFIELD, R.J.W. (eds). Analysis of Visual Behavior. Pp.434. ISBN 0-262-09022-8. (MIT Press: 1982.) $75.

JAKOBY, W.B. (ed.). Methods in Enzymology. Vol.77, Detoxication and Drug Metabolism: Conjugation and Related Systems. Pp.476. ISBN 0-12-181977-9. (Academic: 1981.) $45.

JECHT, E. W. and ZEITLER, E. (eds). Varicocele and Male Infertility. Recent Advances in Diagnosis and Therapy. Pp.211. ISBN 3-540-10727-4/0-387-10727-4. (Springer-Verlag: 1982.) DM88, $45.70.

KIRPICHNIKOV, V.S. Genetic Bases of Fish Selection. Translated from the Russian by G.G. Gause. Pp.410. ISBN 3-540-10911-0. (Springer-Verlag: 1981.) DM 98, $45.70.

KRIEGER, D.T. Cushing's Synodrome, Vol.22. Pp.142. ISBN 3-540-10811-4/0-387-10811-4. (Springer-Verlag: 1982.) DM88, $41.

LAIDLER, E. Otters in Britain. Pp.200. ISBN 0-7153-8069-9. (David & Charles: 1982.) £7.95.

LANGER, R. H. M. and HILL, G. D. Agricultural Plants. Pp.310. Hbk ISBN 0-521-22450-0; pbk ISBN 0-521-29506-8. (Cambridge University Press: 1982.) Hbk £20; pbk £7.95.

LARWOOD, G.P. and NIELSEN, C. Recent and Fossil Bryozoa. Papers Presented at the 5th International Conference, Durham 1980. Pp.334. Pbk ISBN 87-85215-11-2. (Olsen & Olsen, Fredensborg, Denmark: 1981.) Pbk $55.

LESSELL, S. and VAN DALEN, J. T. W. Neuro-ophthalmology 1982, Vol. 2. Pp.404. ISBN 90-219-3059-5/0-444-90210-4. (Excerpta Medica/Elsevier: 1982.) Dfl.190, $88.25.

LEWIS, D. H. (ed.). Induced Skeletal Muscle Ischemia in Man. Symposium held November 6-7, 1980, Linkoping. Pp.180. ISBN 3-8055-3427-2. (Karger, Basel: 1982.) SwFr.115, DM138, $69.

LOTE, C. J. Principles of Renal Physiology. Pp.179. Hbk ISBN 0-7099-0078-3; pbk ISBN 0-0799-0079-1. (Croom Helm: 1982.) Hbk £12.95; pbk £6.95.

THE MACDONALD ENCYCLOPEDIA OF TREES. Pp.300. Pbk ISBN 0-356-08574-0. (Macdonald, London: 1982.) Pbk £4.95.

MARAMOROSCH, K. and RAYCHAUDHURI, S.P. (eds). Mycoplasma Diseases of Trees and Shrubs. Pp.362. ISBN 0-12-470220-1. (Academic: 1981.) $29.50.

MAYERSDORF, A. and SCHMIDT, R.P. (eds). Secondary Epileptogenesis. Pp.188. ISBN 0-89004-578-X. (Raven: 1982.) $21.

MÉNACHÉ, D., SURGENOR, D. MacN. and ANDERSON, H.D. (eds). Hemophilia and Hemostatis. American Red Cross 13th Annual Scientific Symposium, Washington, DC, May 1981. Progress in Clinical and Biological Research, Vol.72. Pp.282. ISBN 0-8451-0072-9. (Alan R. Liss, New York: 1981.) £20.50, DM 90.

MERRELL, D.J. Ecological Genetics. Pp.500. ISBN 0-582-46349-1. (Longman: 1982.) £15.

MITCHELL, A. and WILKINSON, J. The Trees of Britain and Europe. Pp.264. Hbk ISBN 0-00-219037-0; pbk ISBN 0-00-219035-4. (Collins: 1982.) Hbk £6.95; pbk £3.95.

MOYLE, P.B. and CECH, J.J. Jr. Fishes: An Introduction to Ichthyology. Pp.293. ISBN 0-13-319723-9. (Prentice-Hall: 1982.) $29.95.

MYANT, M.B. The Biology of Cholesterol and Related Steroids. Pp.910. ISBN 0-433-22880-6. (Heinemann, London: 1981.) £45.

NATORI, S., IKEKAWA, N. and SUZUKI, M. (eds). Advances in Natural Products Chemistry. Extraction and Isolation of Biologically Active Compounds. Pp.599. ISBN 0-470-27245-7. (Wiley: 1981.) £66, $120.

NESTER, E.W. *et al.* The Microbial Perspective. Pp.730. ISBN 0-03-047041-2. (Saunders College Publishing: 1982.) Np.

NOBILE, S. and WOODHILL, J.M. Vitamin C. The Mysterious Redox-System. A Trigger of Life? Pp.185. ISBN 0-85200-419-2. (MTP Press, Lancaster: 1981.) £11.95.

ODDIE, B. Bill Oddie's Little Black Bird Book. Pp.148. Pbk ISBN 0-413-49480-2. (Methuen: 1982.) Pbk £1.50.

ORR, R.T. Vertebrate Biology. 5th Edn. Pp.568. ISBN 0-03-057959-7. (Saunders: 1982.) Np.

ORTNER, D.J. and PUTSCHAR, G.J. Identification of Pathological Conditions in Human Skeletal Remains. Smithsonian Contributions to Anthropology, No.28. Pp.479. (no ISBN) (Smithsonian Institution: 1981.) Pbk np.

PARKER, T.A. III, ALLEN PARKER, S. and PLENGE, M.A. An Annotated Checklist of Peruvian Birds. Pp.104. ISBN 0-931130-07-7. (Buteo: 1982.) $15.

PATTERSON, B. and WOOD, A.E. Rodents from the Deseadan Oligocene of Bolivia and the Relationships of the Caviomorpha. Bulletin of the Museum of Comparative Zoology, Vol.149: No.7. 12 January 1982. Pp.172. Pbk ISBN 0027-4100. (Harvard University: 1982.) Pbk $3.25.

PENNY, N.D. and ARIAS, J.R. Insects of an Amazon Forest. Pp.269. ISBN 0-231-05266-9. (Columbia University Press: 1982.) $48.75.

PESTKA, S. (ed.). Methods in Enzymology. Vol. 79. Interferons, Pt B. Pp.677. ISBN 0-12-181979-5. (Academic: 1981.) $59.

PICK, E. (ed.). Lymphokines. A Forum for Immunoregulatory Cell Products, Vol. 3. Pp.450. ISBN 0-12-432003-1. (Academic: 1981.) Np.

PODESTA, R. B. (ed.). Membrane Physiology of Invertebrates. Pp.664. ISBN 0-8247-1503-9. (Marcel Dekker: 1982.) Np.

POISSONET, P. *et al.* (eds). Vegetation Dynamics in Grasslands, Heathlands and Mediterranean Ligneous Formations. Advances in Vegetation Science, 4. Symposium of the Working Groups for Succession Research on Permanent Plots, and Data-processing in Phytosociology held at Montpellier, France, September 1980. Pp.286. ISBN 90-6193-636-5. (W. Junk, The Hague: 1982.) Dfl.195, $85.

POMEROY, L.R. and WIEGERT, R.G. (eds). The Ecology of a Salt Marsh. Ecological Studies — Analysis and Synthesis, Vol.38. Pp.271. ISBN 3-540-90555-3. (Springer-Verlag: 1981.) DM 69, $32.20.

RANSOM, R. (ed.). A Handbook of *Drosophila* Development. Pp.289. Hbk ISBN 0-444-80366-1; pbk ISBN 0-444-80418-8. (Elsevier: 1982.) Hbk Dfl.250, $116; pbk Dfl.100, $46.

RATTAZZI, M.C., SCANDALIOS, J.G. and WHITE, G.S. (eds). Isozymes. Current Topics in Biological and Medical Research, Vol.5. Pp.240. ISBN 0-8451-0254-0. (Alan R. Liss, New York: 1981.) £24.60, DM 108.

RAY, D.S. (ed.) The Initiation of DNA Replication. ICN–UCLA Symposia on Molecular and Cellular Biology, Vol.22, 1981. Proceedings of the 1981 ICN–UCLA Symposia held at Salt Lake City, Utah, March, 1981. Pp.628. ISBN 0-12-583580-9. (Academic: 1981.) $53.

RISSER, P.G. *et al.* The True Prairie Ecosystem. Pp.557. ISBN 0-87933-361-8. (Academic: 1981.) $31.50.

ROLLAND, R. *et al.* (eds). Follicular Maturation and Ovulation. Proceedings of the 4th Reinier de Graaf Symposium, August 20-22, 1981, Nijmegen. Pp.432. ISBN 90-219-0505-1/0-444-90231-7. (Excerpta Medica/Elsevier: 1982.) Dfl.195, $90.75.

RUBIN, R.H. and YOUNG, L.S. (eds). Clinical Approach to Infection in the Compromised Host. Pp.653. ISBN 0-306-40679-9. (Plenum: 1981.) $59.50.

RUTHERFORD, G.K. (ed.). The Physical Environment of the Faeroe Islands. Monographiae Biologicae, Vol.46. Pp.147. ISBN 90-6193-099-5. (W. Junk, The Hague: 1982.) Dfl.95, $39.50.

SANDERS, G. L. *et al.* (eds). Pulmonary Toxicology of Respirable Particles. Proc. of the 19th Annual Hanford Life Sciences Symposium, October 22-24, 1979, Richland. Pp.676. ISBN 0-87079-121-4. (US Department of Energy: 1980.) $25.25.

© 1982 Nature Publishing Group

SAN PIETRO, A. (ed.). Biosaline Research: A Look to the Future. Environmental Science Research, Vol.23. Proceedings of the 2nd International Workshop held in La Paz, Mexico, November 1980. Pp.578. ISBN 0-306-40892-9. (Plenum: 1982.) $65.

SATO, G. (ed.). Functionally Differentiated Cell Lines. Pp.280. ISBN 0-8451-0241-1. (Alan R. Liss, New York: 1981.) £25.90, DM 114.

SEGRE, D. and SMITH, L. (eds). Immunological Aspects of Aging. Immunology Series, Vol.15. Pp.584. ISBN 0-8247-1349-4. (Marcel Dekker: 1981.) SwFr. 178.

SILVERSTONE, T. (ed.). Drugs and Appetite. Pp.200. ISBN 0-12-643780-7. (Academic: 1982.) £14.20, $29.50.

SMITH, H. (ed.). Plant and the Daylight Spectrum. Pp.508. ISBN 0-12-650980-8. (Academic: 1981.) £22, $45.50.

SPENCE, A. P. Basic Human Anatomy. Pp.704. ISBN 0-8053-6994-5. (Benjamin/Cummings: 1982.) $27.95.

STAMATOYANNOPOULOS, G. and NIENHUIS, A.W. (eds). Hemoglobins in Development and Differentiation. Pp.526. ISBN 0-8451-0213-3. (Alan R. Liss, New York: 1981.) Np.

TU, A. T. (ed.). Survey of Contemporary Toxicology, Vol. 2. Pp.248. ISBN 0-471-06352-5. (Wiley: 1982.) Np.

WEXLER, P. Information Resources in Toxicology. Pp.291. ISBN 0-444-00616-8. (Elsevier/North-Holland: 1982.) Np.

WOOLHOUSE, H. W. (ed.). Advances in Botanical Research, Vol. 9. Pp.279. ISBN 0-12-005909-6. (Academic: 1981.) £24, $49.50.

Applied Biological Sciences

ANDREOLI, M., MONACO, F. and ROBBINS, J. (eds). Advances in Thyroid Neoplasmia 1981. Pp.363. ISBN 88-7084-000-X. (Acta Medica: 1981.) L10,000.

BARNETT, H. J. M., HIRSH, J. and FRASER MUSTARD, J. (eds). Acetylsalicylic Acid: New Uses for an Old Drug. Pp.288. ISBN 0-89004-647-6. (Raven: 1982.) $53.04.

BLUME, W. T. Atlas of Pediatrics Electroencephalography. Pp.344. ISBN 0-89004-564-X. (Raven: 1982.) $85.

BOZLER, G. and VAN ROSSUM, J. M. (eds). Pharmacokinetics During Drug Development: Data Analysis and Evaluation Techniques. Drug Development and Evaluation, Vol.6. Pp.328. ISBN 3-437-10654-6. (Gustav Fischer: 1982.) DM 98.

CARPENTER, D. C. (ed.). Cellular Pacemakers. Mechanisms of Pacemaker Generation, Vol. 1. Pp.332. ISBN 0-471-06509-9. (Wiley: 1982.) Np.

DE SERRES, F. J. and HOLLAENDER, A. (eds). Chemical Mutagens. Principles and Methods for their Detection, Vol. 7. Pp.479. ISBN 0-306-40771-X. (Plenum: 1982.) $49.50.

FIECHTER, A. (ed.). Advances in Biochemical Engineering. Vol. 21, Microbes and Engineering Aspects. Pp.232. ISBN 3-540-11019-4. (Springer-Verlag: 1982.) DM92, $42.90.

GEORGIEVSKII, V.I., ANNANKOV and SAMOKHIN, V.T. Mineral Nutrition of Animals. Studies in the Agricultural and Food Sciences. Pp.475. ISBN 0-408-10770-7. (Butterworths: 1981.) £45.

HÄMMERLING, G. J., HÄMMERLING, U. and KEARNEY, J. F. (eds). Monoclonal Antibodies and T-Cell Hybridomas: Perspectives and Technical Advances. Research Monographs in Immunology, Vol. 3. Pp.587. ISBN 0-444-80351-3. (Elsevier Biomedical: 1981.) $99.50, Dfl.234.

HAY, J. B. (ed.). Animal Models of Immunological Processes. Pp.260. ISBN 0-12-333520. (Academic: 1982.) £21.20, $43.50.

HOFSCHNEIDER, P. H. and GOEBEL, W. (eds). Gene Cloning in Organisms Other Than *E. coli*. Current Topics in Microbiology and Immunology, Vol.96. Pp.264. ISBN 3-540-11117-4. (Springer-Verlag: 1982.) DM92, $42.90.

HOLLAENDER, A. *et al.* (eds). Genetic Engineering of Microorganisms for Chemicals. Basic Life Sciences, Vol.19. Proceedings of a Symposium held May 1981, at the University of Illinois at Champaign-Urbana. Pp.485. ISBN 0-306-40912-7. (Plenum: 1982.) $55.

HOSKING, G. An Introduction to Paediatric Neurology. Pp.252. Hbk ISBN 0-571-11848-8; pbk ISBN 0-571-11849-6. (Faber & Faber, London: 1982.) Hbk £11.50; pbk £5.95.

HSU, J. M. and DAVIS, R. L. (eds). Handbook of Geriatric Nutrition: Principles and Applications for Nutrition and Diet in Aging. Pp.372. ISBN 0-8155-0880-8. (Noyes, New Jersey: 1981.) $39.

HSU, T. C. (ed.). Cytogenetic Assays of Environmental Mutagens. Pp.412. ISBN 0-916672-56-5. (Allanheld Osmun: 1982.) $45.

LOWRANCE, W. W. (ed.). Assessment of Health Effects at Chemical Disposal Sites. Proceedings of a Symposium held June 1–2, 1981, New York City. Pp.162. ISBN 0-86576-025-X. (William Kaufmann: 1981.) $10.

RHODES-ROBERTS, M. and SKINNER, F. A. (eds). Bacteria and Plants. Society for Applied Bacteriology Symposium, 10. Pp.264. ISBN 0-12-587080-9. (Academic: 1982.) £14.50, $29.50.

Psychology

BRANDSTATTER, H., DAVIS, J. H. and STOCKER-KREICHGAUER, G. (eds). Group Decision Making. European Monographs in Social Psychology, 25. Pp.500. ISBN 0-12-125820-3. (Academic: 1982.) £21.80, $45.

HOFFMEISTER, F. and STILLE, G. (eds). Psychotropic Agents. Part 11, Anxiolytics, Gerontopsychopharmacological Agents, and Psychomotor Stimulants. Pp.778. ISBN 3-540-10300-7/0-387-10300-7. (Springer-Verlag: 1981.) DM 360, $134.10.

KEMPTON, W. The Folk Classification of Ceramics. A Study of Cognitive Prototypes. Pp. 237. ISBN 0-12-404080.2. (Academic: 1981.) $24.50.

MAIN, C. J. (ed.). Clinical Psychology and Medicine: A Behavioural Perspective. Pp.361. ISBN 0-306-40900-3. (Plenum: 1982.) $39.50.

SAHAKIAN, W. S. (ed.). History and Systems of Social Psychology, 2nd Edn. Pp.656. Hbk ISBN 0-89116-179-1; pbk ISBN 0-89116-191-0. (Hemisphere: 1982.) Hbk $24.50, pbk $16.95.

SHEPHERD, M. *et al.* Psychiatric Illness in General Practice, 2nd Edn. Pp.238. ISBN 0-19-261243-3. (Oxford University Press: 1981.) £15.

Anthropology

BAILLIE, M.G.L. Tree-Ring Dating and Archaeology. Pp.274. ISBN 0-7099-0613-7. (Croom Helm, London: 1982.) £16.95.

BROTHWELL, D.R. Digging up Bones: The Excavation, Treatment and Study of Human Skeletal Remains. 3rd Edn. Pp.208. Hbk ISBN 0-19-858504-7; pbk ISBN 0-19-858510-1. (British Museum (Natural History)/Oxford University Press: 1982.) Hbk np; pbk £8.95.

JONES, G.D. and KAUTZ, R.R. (eds). The Transition to Statehood in the New World: New Directions in Archaeology. Pp.254. ISBN 0-521-24075-1. (Cambridge University Press: 1981.) £17.50.

MAI, L.L., SHANKLIN, E. and SUSSMAN, R.W. (eds). The Perception of Evolution. Essays Honoring Joseph B. Birdsell. Anthropology UCLA, Vol. 7, Numbers 1 & 2, 1981. Pp.273. (no ISBN) (University of California, Los Angeles: 1981.) Pbk np.

General

ABDULLAH, T.A. and ZEIDENSTEIN, S.A. Village Women of Bangladesh: Prospects for Change. Women in Development Series, Vol.4. Pp.246. ISBN 0-08-026795-5. (Pergamon: 1981.) £10, $23.

ALLAN, P. and JOLLEY, M. (eds). Nursing, Midwifery & Health Visiting Since 1900. Pp.316. Hbk ISBN 0-571-11839-9; pbk ISBN 0-571-11840-2. (Faber & Faber, London: 1982.) Hbk £7.95; pbk £4.95.

ANDORKA, R. Determinants of Fertility in Advanced Societies. Pp.431. Pbk ISBN 0-416-67350-3. (Methuen: 1982.) Pbk £5.95.

BARNABY, F. and THOMAS, G. The Nuclear Arms Race–Control or Catastrophe? Pp.250. ISBN 0-86187-229-0. (Francis Pinter (Publishers) Ltd., 5 Dryden Street, London WC2E 9NW: 1982.) £12.50.

BARNES, B. T. S. Kuhn and Social Science. Pp.126. Hbk ISBN 0-333-28936-6; pbk ISBN 0-333-28937-4. (Macmillan, London: 1982.) Hbk £12; pbk £3.95.

BERRY, G. A Tale of Two Lakes. The Fight to Save Ennerdale Water and Wastwater. Pp.96. Pbk ISBN 0-9504629-2-6. (Friends of the Lake District, Kendal, Cumbria: 1982.) Pbk £2.25.

BHAT, D.H.S. The Sixth Sense. Pp.105. (no ISBN) (Bharati Bhat, India: 1982.) Rs 30, $5.

BLACK, M. and REED, J. S. (eds). Perspectives on the American South. Annual Review of Society, Politics and Culture, Vol. 1. Pp.409. ISBN 0-677-16260-X. (Gordon & Breach: 1981.) $35.

BLEECKER, D. Gauge Theory and Variational Principles: Global Analysis. Pp.180. ISBN 0-201-10096-7. (Addison-Wesley: 1981.) $17.50.

BOYDEN, S., MILLAR, S., NEWCOMBE, K. and O'NEILL, B. The Ecology of a City and its People: The Case of Hong Kong. Pp.427. ISBN 0-7081-1095-9. (Australian National University Press: 1981.) $A17.50.

BRINKMANN, K. and SCHAEFER, H. (eds). Der Klektrounfall. Pp.324. ISBN 3-540-11003-8/0-387-11003-8. (Springer-Verlag: 1982.) DM 128, $59.60.

COX, R.A.F. (ed.). Offshore Medicine: Medical Care of Employees in the Offshore Oil Industry. Pp.199. ISBN 3-540-11111-5; 0-387-11111-5. (Springer-Verlag: 1982.) DM 80, £19.

FINDLEY, D.F. (ed.). Applied Time Series II. Proceedings of the 2nd Symposium held in Tulsa, Oklahoma, March 1980. Pp.798. ISBN 0-12-256420-0. (Academic: 1981.) $49.50.

FISHER, A.C. Resource and Environmental Economics. Pp.284. Hbk ISBN 0-521-24306-8; pbk ISBN 0-521-28594-1. (Cambridge University Press: 1982.) Hbk £20; pbk £6.95.

INSTITUTE OF SOIL SCIENCE, ACADEMIA SINICA. (eds). Proceedings of Symposium on Paddy Soils, Beijing, The People's Republic of China. Pp.864. ISBN 3-540-10900-5. (Springer-Verlag: 1981.) DM 98, $45.70.

KERR, F.W.L. The Pain Book. Pp.192. Hbk ISBN 0-13-647826-3; pbk ISBN 0-13-647818-2. (Prentice-Hall: 1981.) Np.

LAUDAN, L. Science and Hypothesis: Historical Essays on Scientific Methodology. Pp. 240. Hbk ISBN 90-277-1315-4; pbk ISBN 90-277-1316-2. (Reidel: 1981.) Hbk Dfl. 80, $34.95; pbk Dfl. 37.50, $14.95.

MACKLIN, R. Man, Mind, and Morality. The Ethics of Behavior Control. Pp.130. Pbk ISBN 0-13-551127-5. (Prentice-Hall: 1982.) Pbk £5.65.

MALCOLM, A. and POYSER, J. (eds). Computers and the General Practitioner. Pp.124. ISBN 0-08-026865-X. (Pergamon: 1981.) £10, $20.

MANN, W.B. Was There a Fifth Man? Quintessential Recollections. Pp.170. ISBN 0-08-027445-5. (Pergamon: 1982.) £9.50, $19.50.

NISBET, R.M. and GURNEY, W.S.C. Modelling Fluctuating Populations. Pp.379. ISBN 0-471-28058-5. (Wiley: 1982.) £19.50, $47.50.

OMRAN, A.R. and STANDLEY, C.C. (eds). Family Formation Patterns and Health. Further Studies: An International Collaborative Study in Columbia, Egypt, Pakistan and the Syrian Arab Republic. Pp.464. Pbk ISBN 92-4-156070-3. (WHO: 1981.) Pbk SwFr. 44.

PRYDE, E.H., PRINCEN, L.H. and MUKHERJEE, K.D. (eds). New Sources of Fats and Oils. Symposium at the International Society for Fat Research/American Oil Chemists' World Congress, April/May 1980. Pp.340. (no ISBN) (American Oil Chemists' Society, Champaign, Illinois: 1981.) Np.

RENFREW, C. and WAGSTAFF, M. (eds). An Island Polity. The Archaeology of Exploitation in Melos. Pp.361. ISBN 0-521-23785-8. (Cambridge University Press: 1982.) Np.

STOLLER, B. B. The Birth of the Imagination. Pp.341. (No ISBN) (Stoller Research Co., California: 1982.) Np.

YIN, R. K. Conserving America's Neighborhoods. Pp.195. ISBN 0-306-40795-7. (Plenum: 1982.) $25.

© 1982 Nature Publishing Group